JOURNAL OF MATHEMATICS AND PHYSICS

Massachusetts Institute of Technology

Volume VIII
1928–1929

EDITORS
CLARENCE L. E. MOORE, Professor of Mathematics, Chairman
FREDERICK S. WOODS, Professor of Mathematics
FREDERICK G. KEYES, Professor of Physico-Chemical Research
VANNEVAR BUSH, Professor of Electric Power Transmission
MANUEL S. VALLARTA, Assistant Professor of Physics

MANAGING EDITOR
PHILIP FRANKLIN, Assistant Professor of Mathematics

QA 1
J87
v.8

CONTENTS

	PAGE
Fundamental Table for Lagrangean Coefficients By GEORGE RUTLEDGE	1
Limiting Values of Lagrangean Coefficients By GEORGE RUTLEDGE	13
A Method for Deriving Formulas of Interpolation By PRESCOTT D. CROUT	18
Decomposition of Determinants By LEPINE HALL RICE	56
Principal Directions in a Subspace V_m Immersed in a Riemannian Space V_n By C. L. E. MOORE	65
Hermitian Polynomials and Fourier Analysis By NORBERT WIENER	70
Classes of Functions, Orthogonal on an Infinite Interval, Having the Power of the Continuum By PHILIP FRANKLIN	74
A New Method in the Theory of Quantics. Second Paper By FRANK L. HITCHCOCK	81
The Fourier Transform Theorem By ANDREW C. BERRY	106
A Method for Deriving Formulas for the Approximate Calculation of Integrals By PRESCOTT D. CROUT	119
Equimomental Systems By PHILIP FRANKLIN	129
Hypersurfaces for Which the Codazzi Equations Take the Simplest Form By C. L. E. MOORE	141

363113

Harmonic Analysis and Group Theory 148
By NORBERT WIENER

Note on the Statistical Interpretation of Maxwell's Equations 155
By M. S. VALLARTA

Almost Periodic Functions and Vibrating Systems . . 163
By CARL F. MUCKENHOUPT

An Application of the Invariant Area Properties of Algebraic Polynomials to the Derivation of Formulas for Mechanical Integration 200
By PRESCOTT D. CROUT

A New Tauberian Theorem with Application to the Summability of Fourier Series and Integrals . . . 216
By S. B. LITTAUER

The Operators of Quantum Mechanics 235
By PHILIP FRANKLIN

Geometry of a Complex of Curves in a Riemannian Space of n Dimensions 247
By C. L. E. MOORE

Settlement of Buildings Due to Progressive Consolidation of Individual Strata 266
By CHARLES TERZAGHI

Tests on High-Pressure Pipe Bends 293
By WILLIAM HOVGAARD

FUNDAMENTAL TABLE FOR LAGRANGEAN COEFFICIENTS[1]

By George Rutledge

1. Introduction. The Lagrange polynomial,[2]

$$
\begin{aligned}
&y_0 + \left\{ \frac{A_{11}}{2D}(y_1 - y_{-1}) + \cdots + \frac{nA_{1n}}{2D}(y_n - y_{-n}) \right\} \frac{x}{h} \\
&+ \left\{ \frac{A_{11}}{2D}(y_1 + y_{-1} - 2y_0) + \cdots \right. \\
&\qquad \left. + \frac{A_{1n}}{2D}(y_n + y_{-n} - 2y_0) \right\} \frac{x^2}{h^2} \\
&+ \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (1) \\
&+ \left\{ \frac{A_{n1}}{2D}(y_1 - y_{-1}) + \cdots + \frac{nA_{nn}}{2D}(y_n - y_{-n}) \right\} \frac{x^{2n-1}}{h^{2n-1}} \\
&+ \left\{ \frac{A_{n1}}{2D}(y_1 + y_{-1} - 2y_0) + \cdots \right. \\
&\qquad \left. + \frac{A_{nn}}{2D}(y_n + y_{-n} - 2y_0) \right\} \frac{x^{2n}}{h^{2n}},
\end{aligned}
$$

which takes the values,

$$y_{-n}, \cdots, y_{-2}, y_{-1}, y_0, y_1, y_2, \cdots, y_n, \qquad (2)$$

respectively, when we give to x/h the values,

$$-n, \cdots, -2, -1, 0, 1, 2, \cdots, n, \qquad (3)$$

yields Lagrangean coefficients in the following manner. If in (1) we place $x/h = \lambda$, we obtain an expression linear in the y's,

$$C_{-n}y_{-n} + \cdots + C_{-1}y_{-1} + C_0y_0 + C_1y_1 + \cdots + C_ny_n, \qquad (4)$$

in which the C's are functions of λ, as follows:

[1] Presented to the *American Mathematical Society*, December 27, 1928.

[2] For this form of the Lagrange polynomial and for evaluation of the coefficients see the following papers: This *Journal*, vol. 2 (1922), p. 47, and vol. 6 (1927), p. 240; *Trans. Amer. Math. Soc.*, vol. 26 (1924), p. 113.

$$C_{-j} = \left(\frac{-j}{\lambda} + 1\right)\left\{\frac{A_{1j}}{2D}\lambda^2 + \frac{A_{2j}}{2D}\lambda^4 + \cdots + \frac{A_{nj}}{2D}\lambda^{2n}\right\},$$

$$C_j = \left(\frac{j}{\lambda} + 1\right)\left\{\frac{A_{1j}}{2D}\lambda^2 + \frac{A_{2j}}{2D}\lambda^4 + \cdots + \frac{A_{nj}}{2D}\lambda^{2n}\right\}, \tag{5}$$

$$C_0 = 1 - \sum_{j=1}^{n}\frac{A_{1j}}{D}\lambda^2 - \sum_{j=1}^{n}\frac{A_{2j}}{D}\lambda^4 - \cdots - \sum_{j=1}^{n}\frac{A_{nj}}{2D}\lambda^{2n}.$$

The number A_{ij} is the cofactor of the element in the ith row and jth column of the determinant,

$$D = \begin{vmatrix} 1 & 2^2 & \cdots & n^2 \\ 1 & 2^4 & \cdots & n^4 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 1 & 2^{2n} & \cdots & n^{2n} \end{vmatrix}, \tag{6}$$

and we have [2]

$$\frac{A_{i,j}}{2D} = (-1)^{i+j}\frac{\binom{2n}{n-j}}{(2n)!}\sum^{(i)} r_1^2 r_2^2 \cdots r_{n-i}^2, \tag{7}$$

where $\sum^{(i)} r_1^2 r_2^2 \cdots r_{n-i}^2$ represents the sum of the squares of the $\binom{n-1}{n-i}$ products of the first n integers, excepting j, taken $n - i$ at a time, $i \neq n$. For $i = n$, $\sum^{(i)} r_1^2 r_2^2 \cdots r_{n-i}^2$ is unity.

The expressions (5) for the Lagrangean coefficients are *absolutely convergent for all values of* λ, as n is increased indefinitely. In fact, it follows from the known limit [3] of the sum of the absolute values of the numbers (7), that the absolute values of C_{-j} and of C_j are less than

$$\left(\frac{j}{|\lambda|} + 1\right)\left\{\frac{\pi^2\lambda^2}{3!} + \frac{2\pi^4\lambda^4}{5!} + \frac{3\pi^6\lambda^6}{7!} + \cdots\right\}, \tag{8}$$

and that the absolute value of C_0 is less than

$$1 + 2\left\{\frac{\pi^2\lambda^2}{3!} + \frac{2\pi^4\lambda^4}{5!} + \frac{3\pi^6\lambda^6}{7!} + \cdots\right\}. \tag{9}$$

[3] This *Journal*, vol. 6 (1927), p. 240, (3); p. 241, (6), (7).

In another paper [4] in the current volume of this journal it is shown that the limits of the coefficients, C_0, C_{-j}, C_j are the coefficients respectively of $f(0)$, $f(-j)$, $f(j)$ in the cardinal series of interpolation,

$$\frac{\sin \pi\lambda}{\pi}\left\{\frac{f(0)}{\lambda}+\sum_{j=1}^{\infty}(-1)^j\left(\frac{f(-j)}{\lambda+j}+\frac{f(j)}{\lambda-j}\right)\right\}. \qquad (10)$$

from which it follows that the cardinal series (10) is the limiting *form* of (4), though not always convergent when (4) has a limiting *value*.

The purpose of this paper is to present, both in rational and in decimal form, the numbers (7) for values of n from 1 to 8, inclusive. This tabulation forms a basis from which values of the Lagrangean coefficients (5) may be readily computed for polynomials of any even degree from two to sixteen and for any desired range of λ. The most useful range is, of course, $-1<\lambda<1$.

The table presented in the next section is computed by means of recursion formulas for the sums, $\sum^{Q} r_1^2 r_2^2 \cdots r_{n-1}^2$, involved in (7), as will be explained in a subsequent section. Other formulas provide an exact check, as will also be explained.

2. **Table.** Since the signs of the numbers (7) are given by the factor $(-1)^{i+j}$, it will be sufficient to tabulate absolute values. For the sake of economy in printing, numerators and denominators are printed in separate columns.

[4] This *Journal*, current volume, p. 13.

TABLE OF ABSOLUTE VALUES OF $A_{i,j}/2D$.

i, j, n	Numerator	Den.	Decimal Form (fifteenth decimal retained)
1, 1, 1	1	2!	.50000 00000 00000
1, 1, 2	16	4!	.66666 66666 66666
2, 1, 2	4	4!	.16666 66666 66666
1, 2, 2	1	4!	.04166 66666 66666
2, 2, 2	1	4!	.04166 66666 66666
1, 1, 3	540	6!	.75000 00000 00000
2, 1, 3	195	6!	.27083 33333 33333
3, 1, 3	15	6!	.02083 33333 33333
1, 2, 3	54	6!	.07500 00000 00000
2, 2, 3	60	6!	.08333 33333 33333
3, 2, 3	6	6!	.00833 33333 33333
1, 3, 3	4	6!	.00555 55555 55555
2, 3, 3	5	6!	.00694 44444 44444
3, 3, 3	1	6!	.00138 88888 88888
1, 1, 4	32, 256	8!	.80000 00000 00000
2, 1, 4	13, 664	8!	.33888 88888 88888
3, 1, 4	1, 624	8!	.04027 77777 77777
4, 1, 4	56	8!	.00138 88888 88888
1, 2, 4	4, 032	8!	.10000 00000 00000
2, 2, 4	4, 732	8!	.11736 11111 11111
3, 2, 4	728	8!	.01805 55555 55555
4, 2, 4	28	8!	.00069 44444 44444
1, 3, 4	512	8!	.01269 84126 98412
2, 3, 4	672	8!	.01666 66666 66666
3, 3, 4	168	8!	.00416 66666 66666
4, 3, 4	8	8!	.00019 84126 98412
1, 4, 4	36	8!	.00089 28571 42857
2, 4, 4	49	8!	.00121 52777 77777
3, 4, 4	14	8!	.00034 72222 22222
4, 4, 4	1	8!	.00002 48015 87301
1, 1, 5	3, 024, 000	10!	.83333 33333 33333
2, 1, 5	1, 401, 960	10!	.38634 25925 92592
3, 1, 5	203, 490	10!	.05607 63888 88888
4, 1, 5	11, 340	10!	.00312 50000 00000
5, 1, 5	210	10!	.00005 78703 70370
1, 2, 5	432, 000	10!	.11904 76190 47619
2, 2, 5	524, 280	10!	.14447 75132 27513
3, 2, 5	98, 280	10!	.02708 33333 33333
4, 2, 5	6, 120	10!	.00168 65079 36507
5, 2, 5	120	10!	.00003 30687 83068

TABLE OF ABSOLUTE VALUES OF $A_{i,j}/2D$—(*Continued*).

i, j, n	Numerator	Den.	Decimal Form (fifteenth decimal retained)
1, 3, 5	72, 000	10!	.01984 12698 41269
2, 3, 5	97, 380	10!	.02683 53174 60317
3, 3, 5	27, 405	10!	.00755 20833 33333
4, 3, 5	2, 070	10!	.00057 04365 07936
5, 3, 5	45	10!	.00001 24007 93650
1, 4, 5	9, 000	10!	.00248 01587 30158
2, 4, 5	12, 610	10!	.00347 49779 54144
3, 4, 5	3, 990	10!	.00109 95370 37037
4, 4, 5	390	10!	.00010 74735 44973
5, 4, 5	10	10!	.00000 27557 31922
1, 5, 5	576	10!	.00015 87301 58730
2, 5, 5	820	10!	.00022 59700 17636
3, 5, 5	273	10!	.00007 52314 81481
4, 5, 5	30	10!	.00000 82671 95767
5, 5, 5	1	10!	.00000 02755 73192
1, 1, 6	410, 572, 800	12!	.85714 28571 42857
2, 1, 6	201, 750, 912	12!	.42119 04761 90476
3, 1, 6	32, 915, 520	12!	.06871 69312 16931
4, 1, 6	2, 307, 096	12!	.00481 64682 53968
5, 1, 6	71, 280	12!	.00014 88095 23809
6, 1, 6	792	12!	.00000 16534 39153
1, 2, 6	64, 152, 000	12!	.13392 85714 28571
2, 2, 6	79, 637, 580	12!	.16625 74404 76190
3, 2, 6	16, 757, 235	12!	.03498 36722 88359
4, 2, 6	1, 314, 225	12!	.00274 36755 95238
5, 2, 6	43, 065	12!	.00008 99057 53968
6, 2, 6	495	12!	.00000 10333 99470
1, 3, 6	12, 672, 000	12!	.02645 50264 55026
2, 3, 6	17, 490, 880	12!	.03651 52851 26396
3, 3, 6	5, 299, 360	12!	.01106 33450 91122
4, 3, 6	498, 300	12!	.00104 02888 00705
5, 3, 6	18, 040	12!	.00003 76616 69606
6, 3, 6	220	12!	.00000 04592 88653
1, 4, 6	2, 138, 400	12!	.00446 42857 14285
2, 4, 6	3, 055, 536	12!	.00637 89682 53968
3, 4, 6	1, 031, 250	12!	.00215 29155 64373
4, 4, 6	118, 998	12!	.00024 84292 32804
5, 4, 6	4, 950	12!	.00001 03339 94708
6, 4, 6	66	12!	.00000 01377 86596

TABLE OF ABSOLUTE VALUES OF $A_{i,j}/2D$—(*Continued*).

i, j, n	Numerator	Den.	Decimal Form (fifteenth decimal retained)
1, 5, 6	248, 832	12!	.00051 94805 19480
2, 5, 6	361, 152	12!	.00075 39682 53968
3, 5, 6	127, 776	12!	.00026 67548 50088
4, 5, 6	16, 236	12!	.00003 38955 02645
5, 5, 6	792	12!	.00000 16534 39153
6, 5, 6	12	12!	.00000 00250 52108
1, 6, 6	14, 400	12!	.00003 00625 30062
2, 6, 6	21, 076	12!	.00004 39998 53027
3, 6, 6	7, 645	12!	.00001 59602 80717
4, 6, 6	1, 023	12!	.00000 21356 92239
5, 6, 6	55	12!	.00000 01148 22163
6, 6, 6	1	12!	.00000 00020 87675
1, 1, 7	76, 281, 004, 800	14!	.87500 00000 00000
2, 1, 7	39, 040, 393, 392	14!	.44782 24206 34920
3, 1, 7	6, 880, 401, 528	14!	.07892 33355 37918
4, 1, 7	553, 443, 891	14!	.00634 84140 76278
5, 1, 7	21, 990, 969	14!	.00025 22528 10846
6, 1, 7	417, 417	14!	.00000 47880 84215
7, 1, 7	3, 003	14!	.00000 00344 46649
1, 2, 7	12, 713, 500, 800	14!	.14583 33333 33333
2, 2, 7	16, 041, 857, 832	14!	.18401 20701 05820
3, 2, 7	3, 643, 001, 362	14!	.04178 79418 35684
4, 2, 7	328, 223, 896	14!	.00376 49728 10111
5, 2, 7	13, 849, 836	14!	.00015 88679 45326
6, 2, 7	272, 272	14!	.00000 31231 62845
7, 2, 7	2, 002	14!	.00000 00229 64432
1, 3, 7	2, 825, 222, 400	14!	.03240 74074 07407
2, 3, 7	3, 957, 249, 296	14!	.04539 25999 41211
3, 3, 7	1, 261, 075, 816	14!	.01446 54798 64785
4, 3, 7	135, 208, 073	14!	.00155 09374 08142
5, 3, 7	6, 289, 283	14!	.00007 21427 65285
6, 3, 7	131, 131	14!	.00000 15041 70340
7, 3, 7	1, 001	14!	.00000 00114 82216
1, 4, 7	577, 886, 400	14!	.00662 87878 78787
2, 4, 7	837, 529, 056	14!	.00960 70827 32082
3, 4, 7	295, 539, 244	14!	.00339 00554 82069
4, 4, 7	37, 845, 808	14!	.00043 41196 35508
5, 4, 7	1, 993, 992	14!	.00002 28725 74955
6, 4, 7	45, 136	14!	.00000 05177 43573
7, 4, 7	364	14!	.00000 00041 75351
1, 5, 7	92, 461, 824	14!	.00106 06060 60606
2, 5, 7	136, 085, 040	14!	.00156 09968 73496
3, 5, 7	50, 218, 168	14!	.00057 60398 29513
4, 5, 7	7, 001, 995	14!	.00008 03181 03321
5, 5, 7	417, 417	14!	.00000 47880 84215
6, 5, 7	10, 465	14!	.00000 01200 41352
7, 5, 7	91	14!	.00000 00010 43837

TABLE OF ABSOLUTE VALUES OF $A_{i,j}/2D$—(*Continued*).

i, j, n	Numerator	Den.	Decimal Form (fifteenth decimal retained)
1, 6, 7	9, 878, 400	14!	.00011 33126 13312
2, 6, 7	14, 659, 736	14!	.00016 81581 02185
3, 6, 7	5, 539, 534	14!	.00006 35425 85243
4, 6, 7	808, 808	14!	.00000 92776 30805
5, 6, 7	52, 052	14!	.00000 05970 75249
6, 6, 7	1, 456	14!	.00000 00167 01405
7, 6, 7	14	14!	.00000 00001 60590
1, 7, 7	518, 400	14!	.00000 59464 34517
2, 7, 7	773, 136	14!	.00000 88684 46368
3, 7, 7	296, 296	14!	.00000 33987 36037
4, 7, 7	44, 473	14!	.00000 05101 38468
5, 7, 7	3, 003	14!	.00000 00344 46649
6, 7, 7	91	14!	.00000 00010 43837
7, 7, 7	1	14!	.00000 00000 11470
1, 1, 8	18, 598, 035, 456, 000	16!	.88888 88888 88888
2, 1, 8	9, 809, 014, 026, 240	16!	.46881 96019 14840
3, 1, 8	1, 826, 232, 728, 320	16!	.08728 43792 86694
4, 1, 8	161, 145, 944, 960	16!	.00770 19339 11424
5, 1, 8	7, 469, 965, 360	16!	.00035 70253 01118
6, 1, 8	185, 545, 360	16!	.00000 88680 98422
7, 1, 8	2, 322, 320	16!	.00000 01109 94757
8, 1, 8	11, 440	16!	.00000 00005 46772
1, 2, 8	3, 254, 656, 204, 800	16!	.15555 55555 55555
2, 2, 8	4, 157, 569, 608, 192	16!	.19871 00970 01763
3, 2, 8	996, 775, 780, 000	16!	.04764 06724 59827
4, 2, 8	98, 597, 322, 824	16!	.00471 24366 94713
5, 2, 8	4, 858, 453, 600	16!	.00023 22086 88516
6, 2, 8	125, 100, 976	16!	.00000 59791 72790
7, 2, 8	1, 601, 600	16!	.00000 00765 48108
8, 2, 8	8, 008	16!	.00000 00003 82740
1, 3, 8	789, 007, 564, 800	16!	.03771 04377 10437
2, 3, 8	1, 117, 480, 046, 592	16!	.05340 97055 20816
3, 3, 8	369, 452, 079, 360	16!	.01765 78783 87045
4, 3, 8	43, 262, 803, 584	16!	.00206 77358 90652
5, 3, 8	2, 346, 424, 080	16!	.00011 21468 07025
6, 3, 8	64, 065, 456	16!	.00000 30619 93947
7, 3, 8	851, 760	16!	.00000 00407 09676
8, 3, 8	4, 368	16!	.00000 00002 08767
1, 4, 8	184, 923, 648, 000	16!	.00883 83838 38383
2, 4, 8	270, 898, 729, 920	16!	.01294 75433 90251
3, 4, 8	98, 760, 203, 360	16!	.00472 02215 33010
4, 4, 8	13, 588, 354, 780	16!	.00064 94523 36554
5, 4, 8	827, 306, 480	16!	.00003 95409 25680
6, 4, 8	24, 413, 480	16!	.00000 11668 36742
7, 4, 8	342, 160	16!	.00000 00163 53459
8, 4, 8	1, 820	16!	.00000 00000 86986

TABLE OF ABSOLUTE VALUES OF $A_{i,j}/2D$—(*Continued*).

i, j, n	Numerator	Den.	Decimal Form (fifteenth decimal retained)
1, 5, 8	36, 415, 733, 760	16!	.00174 04817 40481
2, 5, 8	54, 165, 565, 440	16!	.00258 88309 22164
3, 5, 8	20, 615, 678, 720	16!	.00098 53216 91340
4, 5, 8	3, 066, 743, 680	16!	.00014 65743 19028
5, 5, 8	207, 487, 280	16!	.00000 99168 07515
6, 5, 8	6, 690, 320	16!	.00000 03197 62327
7, 5, 8	100, 240	16!	.00000 00047 90948
8, 5, 8	560	16!	.00000 00000 26765
1, 6, 8	5, 419, 008, 000	16!	.00025 90002 59000
2, 6, 8	8, 126, 584, 320	16!	.00038 84082 55471
3, 6, 8	3, 164, 484, 960	16!	.00015 12458 41350
4, 6, 8	491, 170, 680	16!	.00002 34753 91313
5, 6, 8	35, 486, 880	16!	.00000 16960 87385
6, 6, 8	1, 244, 880	16!	.00000 00594 98757
7, 6, 8	20, 160	16!	.00000 00009 63542
8, 6, 8	120	16!	.00000 00000 05735
1, 7, 8	530, 841, 600	16!	.00002 53714 53942
2, 7, 8	799, 985, 664	16!	.00003 82351 33473
3, 7, 8	315, 777, 280	16!	.00001 50925 03518
4, 7, 8	50, 281, 088	16!	.00000 24031 73203
5, 7, 8	3, 786, 640	16!	.00000 01809 81600
6, 7, 8	141, 232	16!	.00000 00067 50151
7, 7, 8	2, 480	16!	.00000 00001 18531
8, 7, 8	16	16!	.00000 00000 00764
1, 8, 8	25, 401, 600	16!	.00000 12140 63714
2, 8, 8	38, 402, 064	16!	.00000 18354 17944
3, 8, 8	15, 291, 640	16!	.00000 07308 60467
4, 8, 8	2,475, 473	16!	.00000 01183 14670
5, 8, 8	191, 620	16!	.00000 00091 58434
6, 8, 8	7, 462	16!	.00000 00003 56644
7, 8, 8	140	16!	.00000 00000 06691
8, 8, 8	1	16!	.00000 00000 00047

3. **Recursion Formulas and Checks.** The basis of the table presented in the preceding section is the tabulation of the numbers, $\sum^{(j)} r_1^2 r_2^2 \cdots r_{n-1}^2$, by means of simple recursion formulas. For $n = 1, 2, 3, 4$, these numbers are readily written down by direct application of the definition, thus:

$n = 1$

1

$n = 2$

2^2	1^2
1	1

$n = 3$

$2^2 3^2$	$1^2 3^2$	$1^2 2^2$
$2^2 + 3^2$	$1^2 + 3^2$	$1^2 + 2^2$
1	1	1

$n = 4$

$2^2 3^2 4^2$		$1^2 3^2 4^2$
$2^2 3^2 + 2^2 4^2 + 3^2 4^2$		$1^2 3^2 + 1^2 4^2 + 3^2 4^2$
$2^2 + 3^2 + 4^2$		$1^2 + 3^2 + 4^2$
1		1
$1^2 2^2 4^2$		$1^2 2^2 3^2$
$1^2 2^2 + 1^2 4^2 + 2^2 4^2$		$1^2 2^2 + 1^2 3^2 + 2^2 3^2$
$1^2 + 2^2 + 4^2$		$1^2 + 2^2 + 3^2$
1		1

(11)

For $n = 4$ the numbers tabulated in (11) are:

	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$i = 1$	576	144	64	36
$i = 2$	244	169	84	49
$i = 3$	29	26	21	14
$i = 4$	1	1	1	1

(12)

If we indicate the value of n involved in the summations by a subscript on the summation sign, it is evident that, for $j < n$,

$$\sum_{n}^{(j)} r_1^2 r_2^2 \cdots r_k^2 = \sum_{n-1}^{(j)} r_1^2 r_2^2 \cdots r_k^2 + n^2 \sum_{n-1}^{(j)} r_1^2 r_2^2 \cdots r_{k-1}^2. \quad (13)$$

Thus in passing from the values (12) to corresponding values for

$n = 5$, the number 26, for example, is replaced by $169 + (25)(26) = 819$. The number 1, $i = 4, j = 2$, is replaced by $26 + (25)(1) = 51$, and similarly for all numbers tabulated in (12).

To obtain the column $j = 5$, $n = 5$, we observe that

$$\sum_{n}^{(n)} r_1^2 r_2^2 \cdots r_k^2 = \sum_{n-1}^{(n-1)} r_1^2 r_2^2 \cdots r_k^2 + (n-1)^2 \sum_{n-1}^{(n-1)} r_1^2 r_2^2 \cdots r_{k-1}^2. \quad (14)$$

Hence for $i = 3, j = 5, n = 5$, for example, we have $49 + (16)(14) = 273$, and for $i = 4, j = 5, n = 5$, similarly, $14 + (16)(1) = 30$. Thus, for $n = 5$, we have:

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
$i=1$	14,400	3,600	1,600	900	576
$i=2$	6,676	4,369	2,164	1,261	820
$i=3$	969	819	609	399	273
$i=4$	54	51	46	39	30
$i=5$	1	1	1	1	1.

(15)

Successive application of formulas (13) and (14) yields corresponding tables for $n = 6, 7, 8$. It should be noted that, for $i = 1$, each of the recursion formulas, (13) and (14), reduces to its second term.

Valuable checks are obtained by means of the sums of the elements in the columns and in the rows of the arrays similar to (12) and (15). First of all it is evident from formula (13) that the sums of elements in columns $j = 1, 2, 3, 4$ of (15) must be 26 times the corresponding sums in (12). That is, for any value of n the sums of elements in columns $j = 1, 2, \cdots, n-1$, are $n^2 + 1$ times the corresponding sums for the array of next lower order. Similarly, from formula (14), the sum of the elements in column $j = n$, is $(n-1)^2 + 1$ times the sum of elements in the last column of the array of next lower order. For $n = 4$, the sums of elements in columns $j = 1, 2, 3, 4$ are respectively 850, 340, 170, 100 and for $n = 5, j = 1, 2, 3, 4, 5$, the corresponding sums are (26)(850), (26)(340), (26)(170), (26)(100), (17)(100).

This serves as a check merely on the mechanical application of the recursion formulas in passing from one array to the next. It would not, in itself, detect an error in the first of the two arrays. For this latter purpose there is available a check of more fundamental character.

It may be observed that in the totality of products involved in

$$\sum_{j=1}^{n} \sum_{n}^{(j)} r_1^2 r_2^2 \cdots r_{n-i}^2, \tag{16}$$

each individual product occurs exactly i times. For, each of the sums $\sum^{(j)} r_1^2 r_2^2 \cdots r_{n-i}^2$ contains $\binom{n-1}{n-i}$ terms; there are $\binom{n}{n-i}$ different terms involved; and each term must appear the same number of times by reason of symmetry. Therefore

$$\sum_{j=1}^{n} \sum_{n}^{(j)} r_1^2 r_2^2 \cdots r_{n-i}^2 = i \sum_{n+1}^{(n+1)} r_1^2 r_2^2 \cdots r_{(i+1)-(i+1)}^2. \tag{17}$$

Formula (17) provides a check which is independent of the recursion formulas. Thus we have, for example,

$$\begin{aligned} 576 + 144 + 64 + 36 &= (1)(820) \\ 244 + 169 + 84 + 49 &= (2)(273) \\ 29 + 26 + 21 + 14 &= (3)\ (30) \\ 1 + 1 + 1 + 1 &= (4)\ \ (1). \end{aligned} \tag{18}$$

It is worthy of remark that formula (17) may be written in the form

$$\sum_{j=1}^{n} \left\{ \frac{1}{j^2} \sum_{n}^{(j)} \frac{1}{r_1^2 r_2^2 \cdots r_{i-1}^2} \right\} = i \sum_{n} \frac{1}{r_1^2 r_2^2 \cdots r_i^2}, \tag{19}$$

by dividing each side of the equality by $(n!)^2$. Then as n becomes infinite we have

$$\begin{aligned} \Sigma \left\{ \frac{1}{j^2} \right\} &= 1 \Sigma \frac{1}{r_1^2} = \frac{\pi^2}{3!} \\ \Sigma \left\{ \frac{1}{j^2} \sum^{(j)} \frac{1}{r_1^2} \right\} &= 2 \Sigma \frac{1}{r_1^2 r_2^2} = 2 \frac{\pi^4}{5!} \\ \Sigma \left\{ \frac{1}{j^2} \sum^{(j)} \frac{1}{r_1^2 r_2^2} \right\} &= 3 \Sigma \frac{1}{r_1^2 r_2^2 r_3^2} = 3 \frac{\pi^6}{7!} \\ &\cdots\cdots\cdots\cdots\cdots, \end{aligned} \tag{20}$$

as previously shown [5] by use of the Bernoullian numbers.

[5] This *Journal*, vol. 6 (1927), p. 247, (32).

4. Lagrangean Coefficients. As an example of the use of the fundamental table we compute the coefficients C_{-1} and C_1 of the interpolation formula (4) for $\lambda = .5, n = 8$. These are two of the seventeen coefficients required for interpolation without differences by means of a polynomial of degree sixteen. From (5),

$$C_{-1} = (-1)\cdot\left\{\frac{A_{11}}{2D}\frac{1}{4} + \frac{A_{21}}{2D}\frac{1}{4^2} + \cdots + \frac{A_{81}}{2D}\frac{1}{4^8}\right\},$$
$$C_1 = (+3)\cdot\left\{\frac{A_{11}}{2D}\frac{1}{4} + \frac{A_{21}}{2D}\frac{1}{4^2} + \cdots + \frac{A_{81}}{2D}\frac{1}{4^8}\right\}. \tag{21}$$

The detailed computation is then as follows:

$$\frac{1}{4}\,(.88888\ 88888\ 88888) = +\ .22222\ 22222\ 22222$$
$$-\frac{1}{4^2}\,(.46881\ 96019\ 14840) = -\ .02930\ 12251\ 19677$$
$$+\frac{1}{4^3}\,(.08728\ 43792\ 86694) = +\ .00136\ 38184\ 26354$$
$$-\frac{1}{4^4}\,(.00770\ 19339\ 11424) = -\ .00003\ 00856\ 79341$$
$$+\frac{1}{4^5}\,(.00035\ 70253\ 01118) = +\ .00000\ 03486\ 57520$$
$$-\frac{1}{4^6}\,(.00000\ 88680\ 98422) = -\ .00000\ 00021\ 65063$$
$$+\frac{1}{4^7}\,(.00000\ 01109\ 94757) = +\ .00000\ 00000\ 06774$$
$$-\frac{1}{4^8}\,(.00000\ 00005\ 46772) = -\ .00000\ 00000\ 00008$$

$$C_{-1} = -\ .19425\ 50763\ 487\cdots$$
$$C_1 = +\ .58276\ 52290\ 463\cdots.$$

Each of the other coefficients C_{-j}, C_j, $j = 2, 3, \cdots, 8$, is obtainable by a similar computation.

LIMITING VALUES OF LAGRANGEAN COEFFICIENTS [1]

By George Rutledge

1. **Introduction.** The Lagrange interpolation formula for the case of $2n$ unit intervals extending from $-n$ to $+n$, has the form

$$\sum_{j=1}^{n} \frac{x(x^2-1^2)\cdots(x^2-\overline{j-1}^2)(x-j)(x^2-\overline{j+1}^2)\cdots(x^2-n^2)}{-j(j^2-1^2)\cdots(j^2-\overline{j-1}^2)(-2j)(j^2-\overline{j+1}^2)\cdots(j^2-n)^2}\, y_{-j}$$

$$+\frac{(x^2-1^2)(x^2-2^2)\quad\cdots\quad(x^2-n^2)}{(-1^2)\quad(-2^2)\quad\cdots\quad(-n^2)}\, y_0$$

$$+\sum_{j=1}^{n}\frac{x(x^2-1^2)\cdots(x^2-\overline{j-1}^2)(x+j)(x^2-\overline{j+1}^2)\cdots(x^2-n^2)}{j(j^2-1^2)\cdots(j^2-\overline{j-1}^2)2j(j^2-\overline{j+1}^2)\cdots(j^2-n^2)}\, y_j, \quad (1)$$

or briefly,

$$\sum_{j=1}^{n} C_{-j}y_{-j} + C_0y_0 + \sum_{j=1}^{n} C_jy_j. \quad (2)$$

It is well-known that

$$\sum_{j=1}^{n} C_{-j} + C_0 + \sum_{j=1}^{n} C_j = 1, \quad (3)$$

identically. The purpose of this note is to derive the limiting form of this identity, and of the formula (1), as n becomes infinite.

It will appear that the limiting form of the identity (3) is the special case of the cardinal series of interpolation (15) of Whittaker [2] in which $f(-j) = f(j) = f(0) = 1$, and that in consequence the limiting form of the Lagrange formula is exactly Whittaker's cardinal series.

2. **Lagrangean Coefficients as Infinite Products.**

It is at once evident that, from (1),

[1] Presented to the *American Mathematical Society*, December 27, 1928.

[2] E. T. Whittaker, Proc. Roy. Soc. Edin., 35 (1915), p. 181; also J. M. Whittaker, Proc. Edin. Math. Soc. (2), 1 (1927), p. 41.

$$C_0 = \frac{(1^2 - x^2)(2^2 - x^2) \cdots (j^2 - x^2) \cdots (n^2 - x^2)}{(n!)^2}$$

$$= \left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right) \cdots \left(1 - \frac{x^2}{n^2}\right), \quad (4)$$

from which it follows that

$$\lim_{n=\infty} C_0 = \frac{\sin \pi x}{\pi x}. \quad (5)$$

In like manner, by use of the identity

$$\frac{(n!)^2}{(-1)^{n-j}2j^2(j^2-1^2)(j^2-2^2)\cdots(j^2-\overline{j-1}^2)(j^2-\overline{j+1}^2)\cdots(j^2-n^2)}$$

$$= \frac{(n!)^2}{\overline{n-j}!\,\overline{n+j}!} = \frac{\binom{2n}{n-j}}{\binom{2n}{n}}, \quad (6)$$

we obtain

$$C_j = (-1)^{n-j}\frac{\binom{2n}{n-j}}{\binom{2n}{n}} \times \frac{x(x^2-1^2)\cdots(x^2-\overline{j-1}^2)(x+j)(x^2-\overline{j+1}^2)\cdots(x^2-n^2)}{(n!)^2}$$

$$= (-1)^{j+1}\frac{\binom{2n}{n-j}}{\binom{2n}{n}}\frac{x}{j-x} \times \frac{(1^2 - x^2)(2^2 - x^2) \cdots (j^2 - x^2) \cdots (n^2 - x^2)}{(n!)^2}$$

$$= (-1)^{j+1}\frac{\binom{2n}{n-j}}{\binom{2n}{n}}\frac{x}{j-x}\left(1-\frac{x^2}{1^2}\right)\left(1-\frac{x^2}{2^2}\right)\cdots\left(1-\frac{x^2}{n^2}\right). \quad (7)$$

From (7) it follows that

$$\lim_{n=\infty} C_j = (-1)^{j+1}\frac{x}{j-x}\frac{\sin \pi x}{\pi x}, \quad (8)$$

and similarly we obtain

$$\lim_{n=\infty} C_{-j} = (-1)^{j+1} \frac{x}{-j-x} \frac{\sin \pi x}{\pi x}. \tag{9}$$

3. **Limiting Form of the Identity.** From (5), (8), and (9), we have

$$\lim_{n=\infty} C_{-n} + \cdots + \lim_{n=\infty} C_0 + \cdots + \lim_{n=\infty} C_n$$

$$= \frac{\sin \pi x}{\pi x} \left\{ (-1)^{n+1} \frac{x}{-n-x} + \cdots + 1 + \cdots + (-1)^{n+1} \frac{x}{n-x} \right\}. \tag{10}$$

The sum (10) may be written

$$\frac{\sin \pi x}{\pi x} \left\{ 1 + \sum_{j=1}^{n} (-1)^{j+1} \frac{2x^2}{j^2 - x^2} \right\}, \tag{11}$$

and since, when x is not an integer,

$$\pi x \csc \pi x = 1 + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{2x^2}{j^2 - x^2}, \tag{12}$$

we obtain, by allowing n to become infinite in (11),

$$\lim_{n=\infty} \{\lim_{n=\infty} C_{-n} + \cdots + \lim_{n=\infty} C_0 + \cdots + \lim_{n=\infty} C_n\}$$

$$= \frac{\sin \pi x}{\pi x} \cdot \frac{\pi x}{\sin \pi x} = 1. \tag{13}$$

Thus the sum,

$$C_0 + (C_1 + C_{-1}) + (C_2 + C_{-2}) + \cdots + (C_n + C_{-n}), \tag{14}$$

becomes, in the limiting case, an absolutely convergent series.

From (10) it follows that the limiting form of the Lagrange formula (1) is the cardinal series

$$\frac{\sin \pi x}{\pi} \left\{ \frac{f(0)}{x} + \sum_{j=1}^{\infty} (-1)^j \left(\frac{f(-j)}{x+j} + \frac{f(j)}{x-j} \right) \right\}, \tag{15}$$

even though the latter may diverge in cases where the former is convergent.

4. **Relation to Expressions for Lagrangean Coefficients in**

Terms of Other Coefficients. The Lagrangean coefficients are expressible in the form [8]

$$C_{-j} = \left(\frac{-j}{\lambda} + 1\right)\left\{\frac{A_{1j}}{2D}\lambda^2 + \frac{A_{2j}}{2D}\lambda^4 + \cdots + \frac{A_{nj}}{2D}\lambda^{2n}\right\},$$

$$C_j = \left(\frac{j}{\lambda} + 1\right)\left\{\frac{A_{1j}}{2D}\lambda^2 + \frac{A_{2j}}{2D}\lambda^4 + \cdots + \frac{A_{nj}}{2D}\lambda^{2n}\right\}, \qquad (16)$$

$$C_0 = 1 - \sum_{j=1}^{n}(C_j + C_{-j}),$$

where $\lambda = \frac{x}{h}$, is used instead of the variable x of the preceding sections in order that the coefficients may apply directly to the interval $(-nh, +nh)$.

Since [9]

$$\frac{A_{ij}}{2D} = (-1)^{i+j}\frac{\binom{2n}{n-j}}{(2n)!}\sum^{\wp} r_1^2 r_2^2 \cdots r_{n-i}^2$$

$$= (-1)^{i+j}\frac{\binom{2n}{n-j}}{\binom{2n}{n}}\frac{1}{j^2}\sum^{\wp}\frac{1}{r_1^2 r_2^2 \cdots r_{i-1}^2}, \qquad (17)$$

where $\sum^{\wp}\frac{1}{r_1^2 r_2^2 \cdots r_{i-1}^2}$ represents the sum of products of squared reciprocals of the first n integers *excepting* j, taken $i-1$ at a time, we have

$$\left\{\frac{A_{1j}}{2D}\lambda^2 + \frac{A_{2j}}{2D}\lambda^4 + \cdots + \frac{A_{nj}}{2D}\lambda^{2n}\right\}$$

$$= (-1)^{j+1}\frac{\binom{2n}{n-j}}{\binom{2n}{n}}\left[\frac{1}{j^2}\lambda^2 - \frac{1}{j^2}\sum^{\wp}\frac{1}{r_1^2}\lambda^4 + \frac{1}{j^2}\sum^{\wp}\frac{1}{r_1^2 r_2^2}\lambda^6 - \cdots\right]$$

[8] This *Journal*, current volume, p. 2.
[9] This *Journal*, loc. cit., p. 2; and vol. 6 (1927), p. 242.

$$= (-1)^{j+1} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} \frac{\lambda^2}{j^2-\lambda^2} \left(1-\frac{\lambda^2}{1^2}\right)\left(1-\frac{\lambda^2}{2^2}\right)\cdots\left(1-\frac{\lambda^2}{n^2}\right). \quad (18)$$

Then from (16) and (18) we have

$$C_{-j} = (-1)^{j+1} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} \frac{\lambda}{-j-\lambda} \left(1-\frac{\lambda^2}{1^2}\right)\cdots\left(1-\frac{\lambda^2}{n^2}\right),$$
$$C_j = (-1)^{j+1} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} \frac{\lambda}{j-\lambda} \left(1-\frac{\lambda^2}{1^2}\right)\cdots\left(1-\frac{\lambda^2}{n^2}\right), \quad (19)$$

in agreement with (7).

From (3), (4), and (19), we have, excluding integral values of λ, the identity

$$\frac{1}{\left(1-\frac{\lambda^2}{1^2}\right)\left(1-\frac{\lambda^2}{2^2}\right)\cdots\left(1-\frac{\lambda^2}{n^2}\right)} = 1 + \sum_{j=1}^{n} (-1)^{j+1} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} \frac{2\lambda^2}{j^2-\lambda^2}, \quad (20)$$

which is the finite form of the relation

$$\frac{\pi\lambda}{\sin \pi\lambda} = \frac{1}{\prod_{j=1}^{\infty}\left(1-\frac{\lambda^2}{j^2}\right)} = 1 + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{2\lambda^2}{j^2-\lambda^2}. \quad (21)$$

Although the derivation of (21) from other infinite developments is readily accessible,[5] it is evident that the identity (20) suffices to establish (21).

[5] Cf. Pierpont, Functions of a Complex Variable, p. 286.

A METHOD FOR DERIVING FORMULAS OF INTERPOLATION *

By Prescott D. Crout.

1. **Introduction.** The formulas which have been derived for the interpolation of values of a function are, for the most part, polynomial approximations of $f(x)$, the given function. These polynomials differ only in the data that is taken to fix them; and they are usually derived either by determining the coefficients of an assumed polynomial in x, as in the case of the formulas of Lagrange,[1] or by making use of the calculus of differences, as in the case of the formulas of Stirling, Gauss, and Bessel.[2] Expressions limiting the error are usually obtained by applying Rolle's Theorem to the error function.[5, 14]

The purpose of this paper is to develop a general method for deriving formulas of interpolation, together with general types of expressions which limit the error; and also to investigate the uniqueness and determinateness of results. Attention will be confined to functions of a single independent variable.

Without loss of generality the origin will be placed where most convenient, and $f(x)$ will be assumed to possess derivatives of such orders and in such intervals as are necessary to give meaning to the expressions which arise.

2. **Formulas of Lagrange.** By way of illustration let us derive a formula of interpolation involving $f_x = f(x\theta)$, $f_{-2} = f(-2h\theta)$, $f_{-1} = f(-h\theta)$, $f_0 = f(0)$, $f_1 = f(h\theta)$, $f_2 = f(2h\theta)$, where $0 \leq \theta \leq 1$, and $h > 0$. Let $\phi(\theta)$ be defined by the expression

$$\phi(\theta) = f_x - K_{-2}f_{-2} - K_{-1}f_{-1} - K_0f_0 - K_1f_1 - K_2f_2 \quad (1)$$

* This paper was prepared under the supervision of Professor George Rutledge, of the Massachusetts Institute of Technology, who offered many valuable suggestions.

[1] *Goursat: Cours d'Analyse Mathématique*, vol. 1, p. 266.
Jordan: Cours d'Analyse, vol. 2, p. 126.

[2] *Radau: Etudes sur les Formules d'Interpolation.*
Thiele: Interpolationsrechnung.
Gibb: Interpolation and Numerical Integration.

where x, h, and the K's are constants. We now differentiate (1) successively with respect to θ, and thus obtain the following relations:

$$\begin{aligned}
\phi^1(\theta) &= xf_x^1 + 2K_{-2}hf_{-2}^1 + K_{-1}hf_{-1}^1 - K_1hf_1^1 - 2K_2hf_2^1 \\
\phi^2(\theta) &= x^2f_x^2 - 4K_{-2}h^2f_{-2}^2 - K_{-1}h^2f_{-1}^2 - K_1h^2f_1^2 - 4K_2h^2f_2^2 \\
\phi^3(\theta) &= x^3f_x^3 + 8K_{-2}h^3f_{-2}^3 + K_{-1}h^3f_{-1}^3 - K_1h^3f_1^3 - 8K_2h^3f_2^3 \\
\phi^4(\theta) &= x^4f_x^4 - 16K_{-2}h^4f_{-2}^4 - K_{-1}h^4f_{-1}^4 - K_1h^4f_1^4 - 16K_2h^4f_2^4
\end{aligned} \tag{2}$$

where f_j^i represents $\dfrac{d^if}{d\theta^i}$ at the point indicated by the subscript j. Letting $\theta \to 0$, noting that when $\theta = 0$

$$f_j^i = f^i(0), \quad i = 0, 1, 2, 3, 4, \quad j = -2, -1, 1, 2, \tag{3}$$

we have the following relations:

$$\begin{aligned}
\phi(0) &= \{1 - K_{-2} - K_{-1} - K_0 - K_1 - K_2\}f(0) \\
\phi^1(0) &= \{x + 2K_{-2}h + K_{-1}h - K_1h - 2K_2h\}f^1(0). \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{4}$$

For a reason to be made clear later, we determine the K's so as to have, for any f,

$$0 = \phi(0) = \phi^1(0) = \phi^2(0) = \phi^3(0) = \phi^4(0); \tag{5}$$

hence from (4) we have

$$\begin{aligned}
K_{-2} + K_{-1} + K_0 + K_1 + K_2 &= 1 \\
-2K_{-2} - K_{-1} + K_1 + 2K_2 &= \frac{x}{h} \\
4K_{-2} + K_{-1} + K_1 + 4K_2 &= \left(\frac{x}{h}\right)^2. \\
&\cdots\cdots\cdots\cdots
\end{aligned} \tag{6}$$

This system of equations is determinate, for its determinant is of the Vandermonde[3] type and hence not zero. Solving this system by the use of determinants, noting that all determinants involved in the solution are of the Vandermonde type, we have

[3] *Pascal: Die Determinanten*, p. 128, p. 197.

immediately

$$K_{-2} = \frac{\left(\frac{x}{h}+1\right)\left(\frac{x}{h}\right)\left(\frac{x}{h}-1\right)\left(\frac{x}{h}-2\right)}{(-2+1)(-2)(-2-1)(-2-2)}$$

$$K_{-1} = \frac{\left(\frac{x}{h}+2\right)\left(\frac{x}{h}\right)\left(\frac{x}{h}-1\right)\left(\frac{x}{h}-2\right)}{(-1+2)(-1)(-1-1)(-1-2)} \qquad (7)$$

$$K_0 = \frac{\left(\frac{x}{h}+2\right)\left(\frac{x}{h}+1\right)\left(\frac{x}{h}-1\right)\left(\frac{x}{h}-2\right)}{(2)(1)(-1)(-2)};$$

. . . . . .

hence the K's are fourth degree polynomials in $\frac{x}{h}$.

Had the determinants involved in the above solution not been of the Vandermonde type, we could not have obtained the results (7) so easily. Since we shall have to deal with more general determinants which are not of this form, it will be well to develop a more general method of simplifying the solution of (6). To this end we place

$$\begin{aligned} K_1 &= U_1 + V_1, & K_2 &= U_2 + V_2, \\ K_{-1} &= U_1 - V_1, & K_{-2} &= U_2 - V_2, \end{aligned} \qquad (8)$$

where the U's and the V's are constants. Obviously this substitution is but a change of notation, due to which the equations (6) assume the following form:

$$\begin{aligned} 2U_1 + 2U_2 + K_0 &= 1 \\ V_1 + 2V_2 &= \frac{1}{2}\left(\frac{x}{h}\right) \\ U_1 + 4U_2 &= \frac{1}{2}\left(\frac{x}{h}\right)^2 \\ V_1 + 8V_2 &= \frac{1}{2}\left(\frac{x}{h}\right)^3 \\ U_1 + 16U_2 &= \frac{1}{2}\left(\frac{x}{h}\right)^4. \end{aligned} \qquad (9)$$

Here it is evident that two equations, which we denote by U, contain the U's alone; two equations, which we denote by V, contain the V's alone; and one equation contains the U's and K_0. The first equation can therefore be neglected, and later used to determine K_0. Let us solve systems U and V for U_1 and V_1 respectively. We have

$$U_1 = \frac{\begin{vmatrix} \frac{1}{2}\left(\frac{x}{h}\right)^2 & 4 \\ \frac{1}{2}\left(\frac{x}{h}\right)^4 & 16 \end{vmatrix}}{\begin{vmatrix} 1 & 4 \\ 1 & 16 \end{vmatrix}}, \qquad V_1 = \frac{\begin{vmatrix} \frac{1}{2}\left(\frac{x}{h}\right) & 2 \\ \frac{1}{2}\left(\frac{x}{h}\right)^3 & 8 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1 & 8 \end{vmatrix}};$$

hence

$$U_1 = \left(\frac{x}{h}\right)\frac{\begin{vmatrix} \frac{1}{2}\left(\frac{x}{h}\right) & 2 \\ \frac{1}{2}\left(\frac{x}{h}\right)^3 & 8 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1 & 8 \end{vmatrix}} = \left(\frac{x}{h}\right)V_1. \tag{10}$$

Similarly

$$U_2 = -\frac{1}{2}\left(\frac{x}{h}\right)V_2;$$

thus by means of the substitution (8) we avoid the use of fourth degree determinants, using one half as many second degree determinants instead. From (8) and (10) we have finally

$$\begin{aligned}
K_{-2} &= \frac{1}{24}\left\{2\left(\frac{x}{h}\right) - \left(\frac{x}{h}\right)^2 - 2\left(\frac{x}{h}\right)^3 + \left(\frac{x}{h}\right)^4\right\} \\
K_{-1} &= \frac{1}{24}\left\{-16\left(\frac{x}{h}\right) + 16\left(\frac{x}{h}\right)^2 + 4\left(\frac{x}{h}\right)^3 - 4\left(\frac{x}{h}\right)^4\right\} \\
K_0 &= \frac{1}{24}\left\{24 - 30\left(\frac{x}{h}\right)^2 + 6\left(\frac{x}{h}\right)^4\right\} \qquad (11) \\
K_1 &= \frac{1}{24}\left\{16\left(\frac{x}{h}\right) + 16\left(\frac{x}{h}\right)^2 - 4\left(\frac{x}{h}\right)^3 - 4\left(\frac{x}{h}\right)^4\right\} \\
K_2 &= \frac{1}{24}\left\{-2\left(\frac{x}{h}\right) - \left(\frac{x}{h}\right)^2 + 2\left(\frac{x}{h}\right)^3 + \left(\frac{x}{h}\right)^4\right\}.
\end{aligned}$$

Let us now consider the expression

$$\phi^i(\theta) = x^i f_x{}^i - (-2)^i K_{-2} h^i f_{-2}{}^i - (-1)^i K_{-1} h^i f_{-1}{}^i - K_1 h^i f_1{}^i - 2^i K_2 h^i f_2{}^i$$

where $i = 1, 2, 3, 4, 5$. Denoting the greatest $|f^i(u)|$ in the interval $-2h \leq u \leq 2h$, or in the interval bounded by the two outermost points if $|x| > 2h$, by $f_m{}^i$, we have

$$|\phi^i(\theta)| \leq h^i f_m{}^i \left\{ \left| \left(\frac{x}{h} \right)^i \right| + |2^i K_{-2}| + |K_{-1}| + |K_1| + |2^i K_2| \right\},$$

or

$$|\phi^i(\theta)| \leq f_m{}^i \sigma_i \tag{12}$$

where

$$\sigma_i = h^i \left\{ \left| \left(\frac{x}{h} \right)^i \right| + 2^i [|K_{-2}| + |K_2|] + [|K_{-1}| + |K_1|] \right\}.$$

Referring to (8), we see that the expression for σ_i will be simplified to some extent when $\frac{x}{h}$ is given numerically; furthermore it is evident from (8) and (10) that σ_i is limited, or that $\sigma_i < M_i$, for $\frac{x}{h}$ is limited and the determinants of systems U and V are not zero. But

$$|\phi^{i-1}(\theta)| = |\phi^{i-1}(\theta) - \phi^{i-1}(0)| = \left| \int_0^\theta \phi^i(\theta) d\theta \right| \leq \int_0^\theta |\phi^i(\theta)| d\theta \leq \theta f_m{}^i \sigma_i \tag{13}$$

since $\phi^{i-1}(0)$ is zero; hence by repeating this process i times we have finally

$$|\phi(\theta)| \leq \frac{\theta^i f_m{}^i \sigma_i}{i!}. \tag{14}$$

Without the relations (5) this process would be impossible; hence the reason for so determining the K's is now evident. We now place $\theta = 1$; then

$$f(x) = K_{-2} f(-2h) + K_{-1} f(-h) + K_0 f(0) + K_1 f(h) + K_2 f(2h) + \epsilon \tag{15}$$

where

$$|\epsilon| \leq \frac{f_m{}^i \sigma_i}{i!} \quad i = 1, 2, 3, 4, 5.$$

We can write $f(x)$ as a polynomial in $\frac{x}{h}$, plus the remainder ϵ, the result being

$$\begin{aligned} f(x) = f(0) &+ \frac{1}{24}\{2f(-2h) - 16f(-h) + 16f(h) - 2f(2h)\}\left(\frac{x}{h}\right) \\ &+ \frac{1}{24}\{-f(-2h) + 16f(-h) - 30f(0) + 16f(h) - f(2h)\}\left(\frac{x}{h}\right)^2 \\ &+ \frac{1}{24}\{-2f(-2h) + 4f(-h) - 4f(h) + 2f(2h)\}\left(\frac{x}{h}\right)^3 \qquad (16) \\ &+ \frac{1}{24}\{f(-2h) - 4f(-h) + 6f(0) - 4f(h) + f(2h)\}\left(\frac{x}{h}\right)^4 \\ &+ \epsilon. \end{aligned}$$

The fact that $f(x) - \epsilon$ is a polynomial of fourth degree in x could be anticipated from (14), since f_m^5 may be a factor; furthermore this interpolation polynomial, which we shall denote by $P_4(x)$, is identical with that derived on the conditions that [4]

$$P(u) = f(u) \qquad u = -2h, -h, 0, h, 2h; \qquad (17)$$

thus we should expect ϵ to vanish at these five points. That this is actually the case is easily verified by writing K_i as the quotient of two determinants; then when x is placed equal to one of the five values, all of the K's but one vanish due to the fact that two columns of the numerators are identical and the denominators are not zero. The K that does not vanish is one, numerically, for the numerator then becomes the determinant of the system and cancels the denominator.

From (12) we see that there is no simple algebraic expression for σ_i, but that there is a different one in each of the six intervals $(-\infty, -2h)$, $(-2h, -h)$, $\cdots$ $(h, 2h)$, $(2h, \infty)$; hence (12) is not a very convenient limit for practical use. In case i is 5, the following method gives a less complicated result. Let $\psi(z)$ be defined by the expression

$$\psi(z) = f(z) - P_4(z) - R(x) \prod_{i=-2}^{i=2} (z - ih) \qquad (18)$$

[4] *Rutledge: Maclaurin Expansion of the Interpolation Polynomial Determined by 2n+1 Evenly Spaced Points, Transactions of the American Mathematical Society*, vol. 26, p. 113; also *Rutledge: The Polynomial Determined by 2n + 1 Points*, this *Journal*, vol. 2, p. 47.

where $R(x)$ is defined by the relation

$$f(x) - P_4(x) = R(x) \prod_{i=-2}^{i=2} (x - ih) = \phi(1) = \epsilon;$$

hence $\psi(z)$ vanishes at the six points $x, -2h, -h, 0, h, 2h$. From Rolle's Theorem we know that $\psi^1(z)$ must have five zeros between the two outermost points, that $\psi^2(z)$ must have four zeros in the same interval, and finally that $\psi^5(z)$ must have at least one zero in this interval. Let $z = \xi$ be this zero; then

$$\psi^5(\xi) = 0 = f^5(\xi) - P_4^5(\xi) - R(x) \left(\frac{d^5}{dz^5} \prod_{i=-2}^{i=2} (z - ih) \right)_{z=\xi}.$$

But

$$P_4^5(\xi) = 0,$$

and

$$\frac{d^5}{dz^5} \prod_{i=-2}^{i=2} (z - ih) = 5!;$$

consequently

$$R(x) = \frac{f^5(\xi)}{5!},$$

and we have finally

$$f(x) = P_4(x) + \frac{f^5(\xi)}{5!} \prod_{i=-2}^{i=2} (x - ih); \tag{19}$$

thus in addition to (15) we also know that

$$\epsilon = \frac{f^5(\xi)}{5!} \prod_{i=-2}^{i=2} (x - ih).$$

We shall now consider the *interpolation formulas based upon values of* $f(x)$ *at* $(2n + 1)$ *points which divide the interval* $(-nh, nh)$ *into* $2n$ *equal subintervals.*[4, 7] Adopting symmetrical notation we have, as in (1),

$$\phi(\theta) = f_\theta - \{K_{-n} f_{-n} + \cdots + K_0 f_0 + \cdots + K_n f_n\} \tag{20}$$

where

$$f_\theta = f(x\theta), \quad f_i = f(ih\theta) \quad i = -n, -n + 1, \cdots 0, \cdots n - 1, n,$$

and $0 \leq \theta \leq 1$. The relation (20) is now differentiated $2n$ times with respect to θ, the resulting system of derivatives being

of the same form as (2). The K's are determined so as to have

$$0 = \phi(0) = \phi^1(0) = \phi^2(0) = \cdots = \phi^{2n}(0) \tag{21}$$

regardless of f, the resulting system of equations being determinate since its determinant is of the Vandermonde type. As in the case of (7), we have immediately [3]

$$K_i = \frac{\prod_{\substack{j=-n \\ j\neq i}}^{j=n} \left(\frac{x}{h} - j\right)}{\prod_{\substack{j=-n \\ j\neq i}}^{j=n} (i - j)}; \tag{22}$$

however, in a more general way as in (8), we place

$$\begin{aligned} K_i &= U_i + V_i \\ K_{-i} &= U_i - V_i \end{aligned} \qquad i = 1, 2, 3, \cdots n. \tag{23}$$

Since

$$K_{-n}(-n)^i + \cdots + K_n(n)^i = \left(\frac{x}{h}\right)^i,$$

we have

$$\begin{aligned} K_0 + 2\sum_{j=1}^{j=n} U_j &= 1, \\ \sum_{j=1}^{j=n} U_j j^i &= \frac{1}{2}\left(\frac{x}{h}\right)^i \quad i = 2, 4, 6, \cdots 2n, \\ \sum_{j=1}^{j=n} V_j j^i &= \frac{1}{2}\left(\frac{x}{h}\right)^i \quad i = 1, 3, 5, \cdots (2n-1). \end{aligned} \tag{24}$$

As in the case of (9) the first equation can be neglected, and later used to determine K_0. We again have two systems of equations,[12] one involving the U's alone, the other the V's alone; furthermore it is evident that

$$U_j = \frac{1}{j}\left(\frac{x}{h}\right) V_j; \tag{25}$$

hence we avoid the use of determinants of degree $2n$, using instead but one half as many of degree n.

To place a limit upon the error function, $\phi(1)$, we write, as in

the case of (12),

$$|\phi^i(\theta)| \leq \left\{ \left| \left(\frac{x}{h} \right)^i \right| + \sum_{j=-n}^{j=n} |K_j j^i| \right\} h^i f_m{}^i$$

$$i = 1, 2, 3, \cdots (2n+1), \quad (26)$$

where $f_m{}^i$ denotes the greatest $|f^i(u)|$ in the interval $-nh \leq u \leq nh$, or in the interval bounded by the two outermost points if $|x| > nh$. Due to the relations (23) it is evident that the expression

$$\sigma_i = \left\{ \left| \left(\frac{x}{h} \right)^i \right| + \sum_{j=-n}^{j=n} |K_j j^i| \right\} h^i$$

can be considerably reduced when $\frac{x}{h}$ is given a definite numerical value. Integrating successively i times as in the case of (12), using (13), we have finally

$$|\phi(\theta)| \leq \frac{\theta^i f_m{}^i \sigma_i}{i!}; \quad (27)$$

hence placing $\theta = 1$ we have

$$f(x) = \{K_{-n} f(-nh) + \cdots + K_0 f(0) + \cdots + K_n f(nh)\} + \epsilon$$

where

$$|\epsilon| \leq \frac{f_m{}^i \sigma_i}{i!} \quad i = 1, 2, 3, \cdots (2n+1). \quad (28)$$

It is evident from (27) that $\sum_{j=-n}^{j=n} K_j f(jh)$ is a polynomial in x of degree $2n$ at most, which we denote by $P_{2n}(x)$; furthermore, reasoning in precisely the same manner as in (17) we have

$$P_{2n}(u) = f(u) \quad u = -nh, \cdots 0, \cdots nh, \quad (29)$$

$(2n+1)$ points in all.

Proceeding as with (18) we now place another limit upon ϵ, and define $\psi(z)$ by the expression [5]

$$\psi(z) = f(z) - P_{2n}(z) - R(x) \prod_{i=-n}^{i=n} (z - ih) \quad (30)$$

[5] *Biermann: Vorlesungen über Mathematische Näherungsmethoden*, p. 97.

where $R(x)$ is defined by the relation

$$f(x) = P_{2n}(x) + R(x) \prod_{i=-n}^{i=n} (x - ih). \tag{31}$$

From (29) and (31) we see that $\psi(z)$ vanishes for z equal to $x, -nh, \cdots nh$, $(2n + 2)$ zeros in all. Applying Rolle's Theorem successively to (30), we finally have $\psi^{2n+1}(\xi) = 0$, where ξ lies somewhere within the interval bounded by the two outermost points; then

$$\psi^{2n+1}(\xi) = f^{2n+1}(\xi) - P_{2n}^{2n+1}(\xi) - R(x)\left(\frac{d^{2n+1}}{dz^{2n+1}} \prod_{i=-n}^{i=n} (z - ih)\right)_{z=\xi}.$$

But

$$P_{2n}^{2n+1}(\xi) = 0,$$

and

$$\frac{d^{2n+1}}{dz^{2n+1}} \prod_{i=-n}^{i=n} (z - ih) = \overline{2n+1}\,!;$$

therefore

$$R(x) = \frac{f^{2n+1}(\xi)}{\overline{2n+1}\,!},$$

and finally from (31) we have

$$f(x) = P_{2n}(x) + \epsilon \tag{32}$$

where [6]

$$\epsilon = \frac{f^{2n+1}(\xi)}{\overline{2n+1}\,!} \prod_{i=-n}^{i=n} (x - ih).$$

The procedure in the case of the *interpolation formulas based upon values of* $f(x)$ *at* $2n$ *points which divide the interval* $\left(-\frac{2n-1}{2}h, \frac{2n-1}{2}h\right)$ *into* $(2n-1)$ *equal subintervals* is identical with that above. The resulting system for determining the K's is

$$\sum_{j=1}^{j=n} K_j \left(\frac{2j-1}{2}\right)^i + \sum_{j=-n}^{j=-1} K_j \left(\frac{2j+1}{2}\right)^i = \left(\frac{x}{h}\right)^i$$

[6] This result can be obtained by different methods.
La Vallée Poussin: Cours d'Analyse Infinitésimale, vol. 2, p. 348.
Nörlund: Leçons sur les Séries d'Interpolation, p. 6.
Nörlund: Differenzenrechnung, p. 10.
Markoff: Differenzenrechnung, p. 6.
Runge und König: Numerisches Rechnen, p. 114.
Steffensen: Interpolation, p. 23.

where $i = 0, 1, 2, \cdots (2n-1)$; hence we have immediately [9]

$$K_i = \frac{\prod_{\substack{j=-n \\ j\neq i}}^{j=-1}\left(\frac{x}{h} - \frac{2j+1}{2}\right) \cdot \prod_{\substack{j=1 \\ j\neq i}}^{j=n}\left(\frac{x}{h} - \frac{2j-1}{2}\right)}{\prod_{\substack{j=-n \\ j\neq i}}^{j=-1}(i-j) \cdot \prod_{\substack{j=1 \\ j\neq i}}^{j=n}(i-j)}.$$

After making the substitution (23), we have [12]

$$\sum_{j=1}^{j=n} U_j\left(\frac{2j-1}{2}\right)^i = \frac{1}{2}\left(\frac{x}{h}\right)^i \quad i = 0, 2, 4, \cdots (2n-2),$$

$$\sum_{j=1}^{j=n} V_j\left(\frac{2j-1}{2}\right)^i = \frac{1}{2}\left(\frac{x}{h}\right)^i \quad i = 1, 3, 5, \cdots (2n-1).$$

Also, reasoning as in the case of (26), we find that

$$|\phi(\theta)| \leq \frac{\theta^i f_m{}^i \sigma_i}{i!} \quad i = 1, 2, 3, \cdots 2n,$$

where

$$\sigma_i = h^i\left\{\left|\left(\frac{x}{h}\right)^i\right| + \sum_{j=1}^{j=n}\left(\left|K_j\left(\frac{2j-1}{2}\right)^i\right| + \left|K_{-j}\left(\frac{-2j+1}{2}\right)^i\right|\right)\right\};$$

hence, placing $\theta = 1$, we have

$$f(x) = \sum_{j=1}^{j=n} K_j f\left(\frac{2j-1}{2}h\right) + \sum_{j=-n}^{j=-1} K_j f\left(\frac{2j+1}{2}h\right) + \epsilon$$

where

$$|\epsilon| \leq \frac{f_m{}^i \sigma_i}{i!} \quad i = 1, 2, 3, \cdots 2n.$$

Furthermore, if we proceed as in the case of (30), we find that

$$\epsilon = \frac{f^{2n}(\xi)}{2n!} \prod_{j=-n+1}^{j=n}\left(x - \frac{2j-1}{2}h\right).$$

We shall finally consider the *formulas of interpolation that are based upon the values of $f(x)$ at n arbitrarily spaced points* $c_1, c_2, \cdots c_n$, where the c's are restricted in neither magnitude nor sign; hence

$$\phi(\theta) = f(x\theta) - \{K_1 f(c_1\theta) + K_2 f(c_2\theta) + \cdots + K_n f(c_n\theta)\} \quad (33)$$

where the c's are known and the K's are to be determined. Placing

$$0 = \phi(0) = \phi^1(0) = \phi^2(0) = \cdots \phi^{n-1}(0)$$

regardless of f, we have a system of equations for determining the K's which has the following determinant:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ c_1 & c_2 & c_3 & \cdots & c_n \\ c_1^2 & c_2^2 & c_3^2 & \cdots & c_n^2 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ c_1^{n-1} & c_2^{n-1} & c_3^{n-1} & \cdots & c_n^{n-1} \end{vmatrix} = \prod_{\substack{i,j=1 \\ i>j}}^{n} (c_i - c_j), \qquad (34)$$

which is of the Vandermonde type and hence not zero, since the points are distinct. Thus the system is determinate, and we have at once [3] that

$$K_i = \frac{\prod_{\substack{j=1 \\ j\neq i}}^{j=n} (x - c_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{j=n} (c_i - c_j)};$$

however, in this case it is useless to make a substitution similar to (23) because of the lack of symmetry of the points.

By proceeding as with (12) we have no trouble in finding that

$$f(x) = \sum_{j=1}^{j=n} K_j f(c_j) + \epsilon \qquad (35)$$

where

$$|\epsilon| \leq \frac{f_m{}^i}{i!} \left\{ |x^i| + \sum_{j=1}^{j=n} |K_j c_j{}^i| \right\} \quad i = 1, 2, 3, \cdots n.$$

It is also to be noted that $\sum_{j=1}^{j=n} K_j f(c_j)$ is a polynomial in x of degree $(n-1)$ at most; furthermore, reasoning as with (17), we find that ϵ vanishes at the n points $c_1, c_2, c_3, \cdots c_n$. Proceeding as with (18) we have finally

$$\epsilon = \frac{f^n(\xi)}{n!} \prod_{i=1}^{i=n} (x - c_i). \qquad (36)$$

3. **Case of Equal Coefficients.**[7] Thus far we have considered the spacing of the points as being initially fixed, the K's representing coefficients to be determined; however, we now alter this point of view. Let

$$\phi(\theta) = f(x\theta) - \{Kf(c_1\theta) + Kf(c_2\theta) + \cdots + Kf(c_n\theta)\} \quad (37)$$

where n points besides x are involved. As before, we differentiate successively and determine the unknown constants by the relations

$$0 = \phi(0) = \phi^1(0) = \cdots \phi^n(0),$$

which give the following system of $(n + 1)$ equations:

$$\begin{aligned} K + K + \cdots + K &= 1 \\ Kc_1 + Kc_2 + \cdots + Kc_n &= x \\ Kc_1^2 + Kc_2^2 + \cdots + Kc_n^2 &= x^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \end{aligned} \quad (38)$$

It is evident that K could not have been fixed arbitrarily, for from the first equation of (38) we have $K = \frac{1}{n}$. A solution of (38) is thus

$$K = \frac{1}{n}, \quad c_1 = c_2 = \cdots c_n = x,$$

a rather surprising result, which is obviously of no practical importance.

4. **Case Where Coefficients and Spacing of Points Are Both Determined.**[7] We shall now leave both the K's and the c's as unknown constants, and place

$$\phi(\theta) = f(x\theta) - \{K_1f(c_1\theta) + K_2f(c_2\theta) + \cdots K_nf(c_n\theta)\}. \quad (39)$$

Differentiating successively, we determine both the K's and the c's so that

$$0 = \phi(0) = \phi^1(0) = \phi^2(0) = \cdots \phi^{2n-1}(0),$$

[7] It may be noted that there is a close correspondence between formulas of interpolation and those of approximate integration. Thus (28) corresponds to the formulas of Cotes, (37) to those of Tchebychef, (39) to those of Gauss, (90) to those of Euler, and so on; in fact the method of derivation of formulas of interpolation that is given in this paper is closely analogous to a method by which formulas of approximate integration are derived.

Crout: A Method for Deriving Formulas for the Approximate Calculation of Integrals, this *Journal*, vol. 7, p. 126.

which process leads directly to the following system of equations:

$$\begin{aligned} K_1 + K_2 + \cdots + K_n &= 1 \\ K_1c_1 + K_2c_2 + \cdots + K_nc_n &= x \\ K_1c_1^2 + K_2c_2^2 + \cdots + K_nc_n^2 &= x^2 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned} \tag{40}$$

A solution of (40) is

$$c_1 = c_2 = c_3 = \cdots = c_n = x$$
$$K_1 + K_2 + K_3 + \cdots + K_n = 1$$

regardless of whether n is even or odd. This result is obviously of no practical value.

5. **Case Where Certain Constants Are Given.** It is evident that certain K's and c's may be fixed, those remaining being then determined in the usual manner. For example, by placing

$$\phi(\theta) = f(x\theta) - \{\tfrac{1}{4}f(0) + K_1f(h\theta) + K_2f(c\theta)\},$$

we find that

$$f(x) = \{\tfrac{1}{4}f(0) + K_1f(h) + K_2f(c)\} + \epsilon$$

where

$$c = \frac{x\left[\frac{x}{h} - 1\right]}{\left[\frac{x}{h} - \frac{3}{4}\right]}, \qquad K_1 = -\frac{\left[\frac{x}{h}\right]^2}{4\left[\frac{x}{h} - \frac{3}{2}\right]\left[\frac{x}{h} - \frac{1}{2}\right]},$$

$$K_2 = \frac{\left[\frac{x}{h} - \frac{3}{4}\right]^2}{\left[\frac{x}{h} - \frac{3}{2}\right]\left[\frac{x}{h} - \frac{1}{2}\right]}, \qquad \epsilon = \frac{f^3(\xi)}{3!}[x][x - h][x - c],$$

or

$$|\epsilon| \leq \frac{f_m{}^i\sigma_i}{i!} \quad i = 1, 2, 3.$$

6. **Formulas Involving Derivatives of $f(x)$.** Thus far the formulas considered have been of the form

$$f(x) = \sum_{i=i_1}^{i=i_2} K_if(c_i) + \epsilon.$$

We shall now consider formulas of the more general type

$$f(x) = \sum_{i=i_{01}}^{i=i_{02}} K_{i,0} f(c_{i,0}) + \sum_{i=i_{11}}^{i=i_{12}} K_{i,1} f^1(c_{i,1}) + \cdots + \epsilon, \quad (41)$$

or

$$f(x) = \sum_{j=0}^{j=j_2} \sum_{i=i_{j,1}}^{i=i_{j,2}} K_{i,j} f^j(c_{i,j}) + \epsilon$$

where $K_{i,j}$ and $c_{i,j}$ are independent of $f(x)$.

As an illustration let us take a specific case, placing

$$\begin{aligned}\phi(\theta) = f_x &- \{A_{-2}f_{-2} + A_2 f_2\} \\ &- \theta\{B_{-2}f_{-2}^1 + B_{-1}f_{-1}^1 + B_1 f_1^1 + B_2 f_2^1\} - \theta^2 C_0 f_0^2 \quad (42)\end{aligned}$$

where $f_x = f(x\theta)$; A_{-2}, A_2, B_{-2}, $\cdots$ B_2, C_0 are independent of f; and the five points in question (those to which the subscripts refer) are such as to divide the interval $(-2h\theta, 2h\theta)$ into four equal subintervals. For convenience in the differentiation that is to come, we note a special case of Leibnitz's rule for the differentiation of a product:

$$\begin{aligned}\psi(\theta) &= K\theta f^1(c\theta) \\ \psi^1(\theta) &= K f^1(c\theta) + Kc\theta f^2(c\theta) \\ \psi^2(\theta) &= 2Kc f^2(c\theta) + Kc^2\theta f^3(c\theta) \quad (43) \\ &\cdots\cdots\cdots\cdots \\ \psi^k(\theta) &= kKc^{k-1}f^k(c\theta) + Kc^k\theta f^{k+1}(c\theta)\end{aligned}$$

where K and c are constants. Using (43) to obtain successive derivatives of $\phi(\theta)$, we have the following relations:

$$\begin{aligned}\phi^1(\theta) &= xf_x^1 + 2A_{-2}hf_{-2}^1 - 2A_2hf_2^1 - B_{-2}f_{-2}^1 \\ &\qquad - B_{-1}f_{-1}^1 - B_1f_1^1 - B_2f_2^1 + \theta\alpha_1 \\ \phi^2(\theta) &= x^2f_x^2 + 4A_{-2}h^2f_{-2}^2 - 4A_2h^2f_2^2 + 4B_{-2}hf_{-2}^2 \\ &\quad + 2B_{-1}hf_{-1}^2 - 2B_1hf_1^2 - 4B_2hf_2^2 - 2C_0f^2(0) + \theta\alpha_2 \\ \phi^3(\theta) &= x^3f_x^3 + 8A_{-2}h^3f_{-2}^3 - 8A_2h^3f_2^3 - 12B_{-2}h^3f_{-2}^3 \\ &\qquad - 3B_{-1}h^3f_{-1}^3 - 3B_1h^3f_1^3 - 12B_2h^3f_2^3 + \theta\alpha_3 \quad (44) \\ \phi^4(\theta) &= x^4f_x^4 - 16A_{-2}h^4f_{-2}^4 - 16A_2h^4f_2^4 + 32B_{-2}h^3f_{-2}^4 \\ &\qquad + 4B_{-1}h^3f_{-1}^4 - 4B_1h^3f_1^4 - 32B_2h^3f_2^4 + \theta\alpha_4 \\ \phi^5(\theta) &= x^5f_x^5 + 32A_{-2}h^5f_{-2}^5 - 32A_2h^5f_2^5 - 80B_{-2}h^4f_{-2}^5 \\ &\qquad - 5B_{-1}h^4f_{-1}^5 - 5B_1h^4f_1^5 - 80B_2h^4f_2^5 + \theta\alpha_5\end{aligned}$$

$$\phi^6(\theta) = x^6 f_x{}^6 - 64A_{-2}h^6 f_{-2}{}^6 - 64A_2 h^6 f_2{}^6 + 192B_{-2}h^5 f_{-2}{}^6 + 6B_{-1}h^5 f_{-1}{}^6 - 6B_1 h^5 f_1{}^6 - 192B_2 h^5 f_2{}^6 + \theta\alpha_6$$

where the $\theta\alpha$'s merely represent terms having θ as a factor. Proceeding as we did in the case of (2), we determine A_{-2}, A_2, B_{-2}, $\cdots$ B_2, C_0 so as to have

$$0 = \phi(0) = \phi^1(0) = \phi^2(0) = \cdots = \phi^6(0) \tag{45}$$

regardless of f. The computations involved in the solution of (45) will be greatly simplified by the following substitution:

$$\begin{aligned} A_2 &= U_2 + V_2, \quad B_2 = R_2 + S_2, \quad B_1 = R_1 + S_1, \\ A_{-2} &= U_2 - V_2, \quad B_{-2} = R_2 - S_2, \quad B_{-1} = R_1 - S_1, \end{aligned} \tag{46}$$

which is similar to (8). We then have

$$\begin{aligned} U_2 &= \tfrac{1}{2} \\ 2V_2h + R_2 + R_1 &= \tfrac{1}{2}x \\ 4U_2h^2 + 4S_2h + 2S_1h + C_0 &= \tfrac{1}{2}x^2 \\ 8V_2h^3 + 12R_2h^2 + 3R_1h^2 &= \tfrac{1}{2}x^3 \\ 16U_2h^4 + 32S_2h^3 + 4S_1h^3 &= \tfrac{1}{2}x^4 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned} \tag{47}$$

In (47) it is quite evident that the adiacent equations have no unknowns in common, whereas the alternate equations have all unknowns in common; hence (47) can be broken up into two separate and smaller systems,[12] as was done with (9). These systems can be solved at once by the use of determinants, it being evident from these solutions and (46) that the A's, B's, and C_0 are polynomials in x of degree six at most. Placing $\theta = 1$ we have

$$f(x) = A_{-2}f(-2h) + A_2 f(2h) + B_{-2}f^1(-2h) + B_{-1}f^1(-h) + B_1 f^1(h) + B_2 f^1(2h) + C_0 f^2(0) + \epsilon. \tag{48}$$

Reasoning as in the case of (12) we can obtain an expression of the form

$$|\epsilon| \leq \frac{f_m{}^i \sigma_i}{i!} + \frac{f_m^{i+1}\tau_i}{i+1!} \quad i = 1, 2, 3, \cdots 7; \tag{49}$$

hence ϵ is zero when f is a polynomial of degree smaller than seven.

In obtaining σ_i and τ_i, all terms involving a particular f_j^i should be collected *before* adding absolute values.

Denoting the right side of (48) less ϵ by $P_6(x)$, we shall now show that

$$\begin{aligned} f(x) &= P_6(x) \quad \text{when} \quad x = -2h, 2h, \\ f^1(x) &= P_6^1(x) \quad \text{when} \quad x = -2h, -h, h, 2h, \\ f^2(x) &= P_6^2(x) \quad \text{when} \quad x = 0. \end{aligned} \tag{50}$$

This is easily done by solving the system of equations resulting from (45) by the use of determinants without making the substitution (46); then each of the quantities $A_{-2}, A_2, \cdots C_0$ is expressed as the quotient of two seventh degree determinants, the denominators being the determinant of the system and different from zero. Since the numerators are obtained by substituting the column whose elements are

$$1, x, x^2, \cdots, \tag{51}$$

successively for the different columns of the determinant of the system, we see that when $x = -2h$ all of the coefficients except A_{-2} vanish due to the fact that two columns of the numerators are identical; furthermore $A_{-2} = 1$, since the numerator is identical with the denominator. Similar relations are true in the case of A_2; hence the validity of the first equation of (50) is proved. We now consider the expression

$$\begin{aligned} P_6^1(x) = {} & A_{-2}^1 f(-2h) + A_2^1 f(2h) + B_{-2}^1 f^1(-2h) \\ & + B_{-1}^1 f^1(-h) + B_1^1 f^1(h) + B_2^1 f^1(2h) + C_0^1 f^2(0), \end{aligned} \tag{52}$$

which is obtained by differentiating $P_6(x)$ with respect to x. The denominators of the coefficients $A_{-2}, A_2, \cdots C_0$ do not contain x and are hence constants in this differentiation. The only part of the numerators which contains x is the column (51); hence we can develop the numerators by minors with respect to this column, differentiate with respect to x, noting that none of the minors contain x, and then build up a new determinant in precisely the same manner in which the other was broken down. We thus see that the derivatives of the coefficients are expressed as quotients of two seventh degree determinants, the denomi-

nators being the determinant of the system and all alike, the numerators being obtained by substituting the column whose elements are

$$\frac{d(1)}{dx}, \frac{d(x)}{dx}, \frac{d(x^2)}{dx}, \frac{d(x^3)}{dx}, \cdots, \tag{53}$$

or,

$$0, 1, 2x, 3x^2, \cdots,$$

successively for the different columns of the determinant of the system; hence when $x = -2h$ all of the coefficients of (52) vanish except B_{-2}^1 due to the fact that two columns of the numerators are identical, as is seen from (44) and (45); furthermore, $B_{-2}^1 = 1$ since the numerator and denominator are identical. Similar relations are true in the case of the other B^1's, the reasoning being the same as with B_{-2}^1; hence the validity of the second equation of (50) is proved. In order to show the truth of the third equation of (50), we consider the expression

$$P_0^2(x) = A_{-2}^2 f(-2h) + A_2^2 f(2h) + B_{-2}^2 f^1(-2h) + B_{-1}^2 f^1(-h) + B_1^2 f^1(h) + B_2^2 f^1(2h) + C_0^2 f^2(0)$$

and reason in the same manner as with (52), no difficulty being encountered. It is easy to see that the relations (50) are of the greatest importance, especially in visualizing the interpolation polynomial $P_0(x)$.

In order to obtain an expression for the error, ϵ, we define $\psi(z)$ by the relation

$$\psi(z) = f(z) - P_0(z) - R(x) \cdot G(z) \tag{54}$$

where $G(z)$ is a seventh degree polynomial such that

$$\begin{aligned} G(u) &= 0 \quad \text{when} \quad u = -2h, 2h, \\ G^1(u) &= 0 \quad \text{when} \quad u = -2h, -h, h, 2h, \end{aligned} \tag{55}$$

and

$$G^2(u) = 0 \quad \text{when} \quad u = 0;$$

and where $R(x)$ is defined by the relation

$$f(x) = P_6(x) + R(x)\cdot G(x). \tag{56}$$

It is thus evident that $\psi(z)$ has at least three zero points: $-2h$, $2h$, and x; hence by Rolle's Theorem $\psi^1(z)$ vanishes at at least two intermediate points. Independent of what has just been said we know from (50) and (55) that $\psi^1(z)$ vanishes at the points $-2h$, $-h$, h, $2h$; hence unless one or more zeros of the two sets of points overlap, we have six points at which $\psi^1(z)$ vanishes. We shall temporarily assume that $f(z)$ is such that there is no overlapping of zeros; then again by Rolle's Theorem we know that $\psi^2(z)$ vanishes at five intermediate points. But as before we know independently from (50) and (55) that $\psi^2(z)$ vanishes at the point 0; hence unless there is overlapping we have six points at which $\psi^2(z)$ vanishes. Applying Rolle's Theorem successively, we know that $\psi^3(z)$ has five intermediate zero points, and so on till finally it is found that $\psi^7(z)$ has at least one intermediate zero point, which we denote by ξ. Then

$$\psi^7(\xi) = 0 = f^7(\xi) - P_6^7(\xi) - R(x)\cdot G^7(\xi),$$

and since $P_6^7(\xi)$ is zero, we have

$$R(x) = \frac{f^7(\xi)}{G^7(\xi)}. \tag{57}$$

We now determine $G(z)$ from the conditions (55); hence placing

$$G(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{n-1}z^{n-1} + z^n,$$

we obtain the following system of equations:

$$\begin{aligned}
&a_0 + a_1 z_1 + a_2 z_1^2 + a_3 z_1^3 + a_4 z_1^4 + a_5 z_1^5 + a_6 z_1^6 = -z_1^7\\
&a_0 + a_1 z_2 + a_2 z_2^2 + a_3 z_2^3 + a_4 z_2^4 + a_5 z_2^5 + a_6 z_2^6 = -z_2^7\\
&a_1 + 2a_2 z_3 + 3a_3 z_3^2 + 4a_4 z_3^3 + 5a_5 z_3^4 + 6a_6 z_3^5 = -7z_3^6 \qquad (58)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_1 + 2a_2 z_6 + 3a_3 z_6^2 + 4a_4 z_6^3 + 5a_5 z_6^4 + 6a_6 z_6^5 = -7z_6^6\\
&\qquad 2a_2 + 6a_3 z_7 + 12a_4 z_7^2 + 20a_5 z_7^3 + 30a_6 z_7^4 = -42z_7^5
\end{aligned}$$

where $z_1 = -2h$, $z_2 = 2h$, $z_3 = -2h$, $\cdots z_6 = 2h$, $z_7 = 0$. It is important to note that the determinant of (58) is identical with

that of (45) when the rows and columns are interchanged; hence it is not zero, and (58) is determinate. Since the coefficient of z^7 in $G(z)$ is one, we have

$$G^7(z) = 7!;$$

hence (56) becomes

$$f(x) = P_6(x) + \epsilon$$

where

$$\epsilon = \frac{f^7(\xi)}{7!} G(x). \tag{59}$$

In the discussion of (56), had $f(x)$ been such that overlapping of zeros occurred, Rolle's Theorem could not be used as above, and *the result* (59) *is not necessarily true.*[15]

7. **General Case of Symmetrical Points.—Balanced Systems.** Let the expression

$$\theta^i[\{K^i_{-n}\}f^i_{-n_i} + \{K^i_{-(n-1)}\}f^i_{-(n_i-1)} + \cdots + \{K^i_{(n-1)}\}f^i_{(n_i-1)} + \{K_n{}^i\}f^i_{n_i}] \tag{60}$$

be known as an element and denoted by $\{E^i\}$, where the subscripts refer to the points

$$\{c^i_{-n}\}\theta, \quad \{c^i_{-(n-1)}\}\theta, \quad \cdots, \quad \{c^i_{(n-1)}\}\theta, \quad \{c_n{}^i\}\theta, \tag{61}$$

which are subject only to the condition that

$$0 \leq -\{c^i_{-j}\} = \{c_j{}^i\} \quad \text{where} \quad j = 1, 2, 3, \cdots n_i, \tag{62}$$

and i is an integer; thus except for symmetry the points are chosen arbitrarily. The $\{c_j{}^i\}$'s and the $\{K_j{}^i\}$'s are independent of θ and f, the upper subscripts merely identifying the element. Also, since there is no interdependence between the points of different elements, we must refer to the points of $\{E^i\}$ by the subscripts

$$1_i, 2_i, 3_i, \cdots n_i;$$

however, where there is no ambiguity superfluous symbols will be omitted; thus we write $\{K_j{}^i\}$ and $\{c_j{}^i\}$ instead of $\{K_{j_i}{}^i\}$ and $\{c_{j_i}{}^i\}$. Also by definition

$$f_{j_i}{}^k = \frac{d^k f}{d\theta^k} \tag{63}$$

at the point $\{c_j{}^i\}\theta$, as in previous sections. We now arbitrarily

construct [8] a certain number of elements and denote $\sum \{E^i\}$ by S. $\phi(\theta)$ is defined by the relation

$$\phi(\theta) = f_x - S = f(x\theta) - \sum \{E^i\} \tag{64}$$

where $f_x = f(x\theta)$ and S may have mid-terms. Let θ^i be taken within the bracket of $\{E^i\}$ so that

$$\{E^i\} = \sum_{j=-n}^{j=n} \theta^i \{K_j{}^i\} f_j{}^i,$$

in which j may or may not take the value zero.

We now take successive derivatives of $\phi(\theta)$, *no attempt being made to collect terms.* Terms not of the form $\theta\alpha$ are called *principal terms*; also, we define [9]

$$\{E_r{}^i\} = \frac{d^r}{d\theta^r} \{E^i\}. \tag{65}$$

Since we are later going to place $\theta = 0$, we need consider only the principal terms of the various derivatives of $\phi(\theta)$. The expression for the principal terms of $\{E_r{}^i\}$ can be found directly from Leibnitz's rule,[10] which is

$$(uv)^r = u^r v + \binom{r}{1} u^{r-1}v^1 + \binom{r}{2} u^{r-2}v^2 + \cdots + \binom{r}{2} u^2 v^{r-2} + \binom{r}{1} u^1 v^{r-1} + uv^r \tag{66}$$

where the upper subscripts refer to derivatives, and

$$\binom{n}{m} = \frac{n!}{n-m! \cdot m!} = \binom{n}{n-m} \quad \text{if} \quad n > m > 0,$$

$$\binom{n}{m} = 0 \quad \text{if} \quad n < m, \tag{67}$$

$$\binom{n}{0} = \binom{n}{n} = 1.$$

[8] The number and choice of elements is arbitrary, as is the number and choice of the points of the different elements. The symmetry of the points is the principal restriction, the other restriction being one that will be discussed later—that the system for determining the K's be determinate.

[9] It is evident that in the above notation $f^0(x) = f(x)$, and $\{E^0\} = \{E_0{}^0\} = \Sigma \{K\} f$.

[10] *Pierpont: Theory of Functions of Real Variables*, vol. 1, p. 253, p. 61.

Thus we can write

$$(uv)^r = \sum_{i=0}^{i=r} \binom{r}{i} u^{r-i} v^i. \tag{68}$$

We now place

$$\{K_j^i\}\theta^i = u,$$
$$f_j^i = v;$$

and note that

$$\begin{aligned} u^k &= 0 \quad \text{if} \quad k > i, \\ u^k &= \theta\alpha \quad \text{if} \quad k < i; \end{aligned} \tag{69}$$

hence, since we are interested only in the principal term, $k = i$ and $u^i = \{K_j^i\}i!$. Thus the expression for the principal term of $\{E_r^i\}$ due to the term $\theta^i\{K_j^i\}f_j^i$ of $\{E^i\}$ is

$$\binom{r}{i}\{K_j^i\}i!\{c_j^i\}^{r-i}f^r(\{c_j^i\}\theta). \tag{70}$$

For convenience, we shall place (70) in a somewhat different form. Let us define

$$\{n, m\} = \overline{m+1}!\binom{n}{m+1} = \frac{n!}{\overline{n-m-1}!};$$

then

$$\{n, m\} = n(n-1)(n-2)\cdots(n-m) \quad \text{if} \quad (n-1) \geq m \geq 0,$$
$$\{n, m\} = 0 \quad \text{if} \quad (n-1) < m,$$
$$\{n, -1\} = 1 \quad \text{for any} \quad n \geq 0.$$

Thus (70) is reduced to the following expression:

$$\{K_j^i\}\{r, i-1\}\{c_j^i\}^{r-i}f^r(\{c_j^i\}\theta) \tag{71}$$

where

$$\{r, i\} = r(r-1)(r-2)\cdots(r-i) \quad \text{if} \quad (r-1) \geq i \geq 0,$$
$$\{r, i\} = 0 \quad \text{if} \quad (r-1) < i,$$
$$\{r, -1\} = 1 \quad \text{for any} \quad r \geq 0.$$

The K's are to be determined so that for any f

$$0 = \phi(0) = \phi^1(0) = \phi^2(0) = \cdots; \tag{72}$$

therefore, as in (44), we place $\theta = 0$. The $\theta\alpha$ terms fall out,

and only principal terms remain; also, when $\theta = 0$

$$f^r{}_{j_i} = f^r(\{c_j{}^i\} \cdot 0) = f^r(0)$$

for any j_i. Since $\frac{d^r f(x\theta)}{d\theta^r} = x^r f_x{}^r$, we see from (64) and (71) that the principal terms of $\phi^r(\theta)$ involve f^r, and no other derivative of f; thus we can factor $f^r(0)$ out of $\phi^r(0)$, as was done in the case of (4), and thus obtain

$$\phi^r(0) = 0 = Q_r f^r(0) \tag{73}$$

where Q_r involves the K's and c's. Since $\phi^r(0)$ is to vanish for any f, we have [11]

$$Q_r = 0 \qquad r = 0, 1, 2, 3, \cdots, \tag{74}$$

as the system for determining the K's.

Let $\{Q_r{}^i\}$ denote the terms of Q_r due to $\{E^i\}$, the one due to $f(x\theta)$ being x^r. The expressions $\{Q_r{}^i\}$ will be known as elements since

$$Q_r = x^r + \sum \{Q_r{}^i\}; \tag{75}$$

thus from (71) we have

$$-\{Q_r{}^i\} = \sum_{j=-n}^{j=n} \{K_j{}^i\}\{r, i-1\}\{c_j{}^i\}^{r-i}. \tag{76}$$

We now make the substitution

$$\begin{aligned} \{K_j{}^i\} &= \{U_j{}^i\} + \{V_j{}^i\} \\ \{K_{-j}{}^i\} &= \{U_j{}^i\} - \{V_j{}^i\} \end{aligned} \qquad j > 0; \tag{77}$$

hence (76) assumes the following form:

$$\begin{aligned} -\{Q_r{}^i\} &= 2\sum_{j=1}^{j=n} \{U_j{}^i\}\{r, i-1\}\{c_j{}^i\}^{r-i} \quad (r-i) \text{ even}, \\ -\{Q_r{}^i\} &= 2\sum_{j=1}^{j=n} \{V_j{}^i\}\{r, i-1\}\{c_j{}^i\}^{r-i} \quad (r-i) \text{ odd}, \end{aligned} \tag{78}$$

where $r \neq i$ or when $r = i$ and $\{E^i\}$ has no mid-term. When $r = i$ and $\{E^i\}$ has a mid-term, the term $\{K_0{}^i\}i!$ must be added

[11] From (73) it is seen that if for a particular $f(x)$ it is known that $f^i(0) = 0$, it is not necessary to place $Q_i = 0$; hence a Q of higher order can be used to determine the K's. In this case the expression limiting the error will involve derivatives of a higher order than usual.

to the expressions (78). It is evident from (78) that the adjacent equations of (74) have no unknowns in common, and that the alternate equations involve the same unknowns; hence (74) can be broken up into two separate and smaller systems of equations, each of which can be solved by the use of determinants.[12]

We shall now show that the expression

$$\epsilon = f(x) - P(x) = f(x) - \sum \{E^i\}_{\theta=1}$$

vanishes at the points $\{c_{-n}{}^0\}, \cdots \{c_n{}^0\}$; and that in general

$$\frac{d^r\epsilon}{dx^r} = f^r(x) - P^r(x) \qquad r = 0, 1, 2, \cdots \tag{79}$$

vanishes at the points $\{c_{-n}{}^r\}, \cdots \{c_n{}^r\}$. To this end let us solve the system (74) by the use of determinants without making the substitution (77), it being assumed that the determinant of the system is different from zero. Then each unknown is expressed as the quotient of two determinants, the denominator of each quotient being the determinant of the system. The numerator is obtained in each case by replacing the column of the determinant of the system that corresponds to the particular unknown by the column whose elements are

$$1, x, x^2, x^3, \cdots. \tag{80}$$

Referring to (64), placing $\theta = 1$, we see that $P^r(x)$ is obtained from $P(x)$ by replacing the K's by the corresponding $\left(\frac{d^rK}{dx^r}\right)$'s. Reasoning as in the case of (52), we see that $\frac{d^rK}{dx^r}$ is obtained from K by replacing the column (80) of the numerator by the column composed of the rth derivatives of the elements of (80). Let us now place $x = \{c_j{}^r\}$, where $j = -n, \cdots n$; then the element of this column lying in the kth row is

$$\left(\frac{d^r}{dx^r}x^{k-1}\right)_{x=\{c_j{}^r\}} = \{k-1, r-1\}\{c_j{}^r\}^{k-r-1}. \tag{81}$$

But the corresponding element of the denominator is found from (71) to be

$$\{k-1, r-1\}\{c_j{}^r\}^{k-r-1},$$

[12] It is obvious that if (74) is determinate, the two systems U and V are determinate, for $\{U_j{}^i\} = \frac{1}{2}[\{K_j{}^i\} + \{K_{-j}{}^i\}]$, and $\{V_j{}^i\} = \frac{1}{2}[\{K_j{}^i\} - \{K_{-j}{}^i\}]$.

which is identical with (81); hence $\{K_f'\}$ is one, since the numerator and denominator are identical; and all of the other K's vanish, due to the fact that two columns of the numerators are alike. The relations (79) are thus necessarily true.

We shall now derive an expression for $P(x)$ of the form $\sum_{i=0}^{i=n} m_i G_i(x)$, which is a *generalization of Newton's representation* and which contains Newton's formulas as special cases. Let $G(x)$ denote a polynomial of lowest degree subject to the conditions

$$G^i(\{c_j{}^i\}) = 0$$

for every point of P; hence placing

$$G(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + x^n,$$

we have

$$\begin{aligned}
&a_0 + a_1x_1 + a_2x_1{}^2 + a_3x_1{}^3 + \cdots + a_{n-1}x_1{}^{n-1} = -x_1{}^n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&a_0 + a_1x_{N_0} + a_2x_{N_0}{}^2 + a_3x_{N_0}{}^3 + \cdots + a_{n-1}x_{N_0}{}^{n-1} = -x_{N_0}{}^n \qquad (82) \\
&a_1 + 2a_2x_{N_0+1} + 3a_3x_{N_0+1}{}^2 + \cdots \\
&\qquad\qquad + (n-1)a_{n-1}x_{N_0+1}{}^{n-2} = -nx_{N_0+1}{}^{n-1} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

where $x_1 = \{c_{-n}{}^0\}$, $\cdots$ $x_{N_0} = \{c_n{}^0\}$, $x_{N_0+1} = \{c_{-n}{}^1\}$, and so on. If we denote the total number of points of P by λ, it is evident that the system (82) contains λ equations; also, if $n = \lambda$ and the rows and the columns of the determinant of (82) be interchanged, the resulting determinant is identical with that of (74). In order to prove this let us consider the element in the column of the determinant of (74) that involves $\{c_j{}^i\}$ and that lies in the kth row. From (71) we see that this element is $\{k-1, i-1\}\{c_j{}^i\}^{k-i-1}$; but the corresponding element of the determinant of (82) is

$$\left(\frac{d^i}{dx^i}x^{k-1}\right)_{x=\{c_j{}^i\}} = \{k-1, i-1\}\{c_j{}^i\}^{k-i-1},$$

which is exactly the same; hence the two determinants are

identical; and since (74) is determinate, (82) is also.[13] We thus have

$$G(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{\lambda-1}x^{\lambda-1} + x^\lambda \qquad (83)$$

where the a's are known. It is to be noted that the coefficient of x^λ is not fixed by (82), and is placed equal to one for convenience alone. It is not necessary to solve (82) in order to obtain $G(x)$, for if we let $f(x) = -x^\lambda$, we obtain an interpolation polynomial which we denote by $P\{x\}$; then it is evident from (79) that

$$G(x) = P\{x\} + x^\lambda. \qquad (84)$$

We now define $G_r(x)$ as a polynomial of degree r in which the coefficient of x^r is one and which is fixed by the conditions

$$G_r{}^i(\{c_j{}^i\}) = 0 \qquad (85)$$

for the first r of the λ points of P; thus $G_\lambda(x) = G(x)$; also, we place $G_0(x) = 1$. Let us now consider the expression

$$\prod(x) = P_{\lambda-1}(x) + m_\lambda G_\lambda(x) \qquad (86)$$

where m_λ is a constant determined by a new condition of the form (79) and the subscript of P merely shows the degree of the polynomial; then $\prod(x)$ evidently satisfies this condition together with those of (79); hence $\prod(x) = P_\lambda(x)$, where $P_\lambda(x)$ is the interpolation polynomial of degree λ that is based upon these $(\lambda + 1)$ conditions; thus *any interpolation polynomial may be taken as the starting point in the derivation of another of higher degree providing that the one of higher degree satisfies the conditions by which the one of lower degree can be fixed.* Applying (86) to $P_{\lambda-1}(x)$ we have

$$P_\lambda(x) = P_{\lambda-2}(x) + m_{\lambda-1}G_{\lambda-1}(x) + m_\lambda G_\lambda(x);$$

and repeating this process, we have finally

$$P_\lambda(x) = m_0G_0(x) + m_1G_1(x) + m_2G_2(x) + \cdots + m_\lambda G_\lambda(x),$$

or

$$P_\lambda(x) = \sum_{i=0}^{i=\lambda} m_iG_i(x). \qquad (87)$$

[13] It is evident that a polynomial can be determined by the relations (79) since the determinant of the resulting system of equations is identical with that of (82) and hence not zero; consequently this polynomial must be identical with $P(x)$ since it is uniquely determined.

The importance of (87) lies in the fact that the m's are invariant as λ is increased. As an example let us consider the formulas of Lagrange (35). In this case $G_j(x) = \prod_{i=1}^{i=j}(x - c_i)$ unless $j = 0$, in which case $G_0(x) = 1$; hence from (87) we have

$$P_{n-1}(x) = m_0 + m_1(x - c_1) + m_2(x - c_1)(x - c_2) + \cdots + m_{n-1}(x - c_1)\cdots(x - c_{n-1}),$$

or

$$P_{n-1}(x) = m_0 + \sum_{j=1}^{j=n-1} m_j \prod_{i=1}^{i=j}(x - c_i),$$

which is Newton's formula for n points.

As in the case of (54) we can consider an expression of the form [14]

$$\psi(z) = f(z) - P(z) - R(x)\cdot G(z)$$

where $R(x)$ is defined by the relation

$$f(x) = P(x) + R(x)\cdot G(x);$$

then

$$\psi(z) = [R(z) - R(x)]G(z).$$

Proceeding as in the case of (56) we obtain an expression of the form

$$\epsilon = \frac{f^{\lambda}(\xi)}{\lambda!} G(x) \tag{88}$$

providing that there is no overlapping of zeros. If overlapping occurs, *such an expression may be false*; [15] in this case nothing can be concluded by the application of Rolle's Theorem. For $\epsilon = \frac{f^{\lambda}(\xi)}{\lambda!} G(x)$, it is necessary and sufficient that $\psi^{\lambda}(\xi) = 0$, for

[14] This is an extension of a method that has been used to obtain an expression for the error when no elements are omitted and when no element has a point not contained in the preceding element, in which case it is evident that there can be no overlapping of zeros.
Biermann: Vorlesungen über Mathematische Näherungsmethoden, p. 101.
Goursat: Cours d'Analyse Mathématique, vol. 1, p. 76.

[15] For example, in the case where

$$f(x) - \epsilon = P_3(x) = [K_{-1}^{0}]f(-1) + [K_1^{0}]f(1) + [K_0^{2}]f^{2}(0),$$

$\epsilon + [f^{3}(\xi)/3!]G(x)$ if f is any polynomial of fourth degree and $|x| < (2 - \sqrt{3})$.

then

$$\psi^{\lambda}(\xi) = f^{\lambda}(\xi) - \frac{f^{\lambda}(\xi)}{\lambda!}\lambda! = 0;$$

therefore, since there is no simple way to tell whether there is or is not overlapping of zeros, it is in every case necessary to make sure that $\psi^{\lambda}(\xi)$ vanishes—a necessity which renders this expression for ϵ worthless in most cases from the practical point of view.

It is evident that we can operate directly with $\phi^{i}(\theta)$ in the same manner as with (49), and thus obtain *an expression limiting the error which is of the form* [16]

$$|\epsilon| \leq \frac{f_m^{\,i}}{i!}\sigma_i + \frac{f_m^{\,i+1}\tau_i}{i+1!} + \cdots + \frac{f_m^{\,i+k}\omega_i}{i+k!} \tag{89}$$

where $\sigma_i, \tau_i, \cdots \omega_i$ are independent of f and $i = 1, 2, 3, \cdots \lambda$; hence P is a polynomial in x of degree $(\lambda - 1)$ at most.

We shall now consider a method by which the error can be limited by a series, finite or infinite, which involves increasing orders of the derivatives of $f(x)$. Let us consider an element of $\phi^{r}(\theta)$, and expand each of the terms of this element about the same point using Taylor's formula, finite or infinite. We then operate upon this new expression in the usual manner, thus obtaining a limit of $|\epsilon|$. Obviously, any or all of the terms of any or all of the elements can be treated in this manner, the particular Taylor's formula used in each case being arbitrary, as is the point about which the expansion is made. In particular, if we expand every term about the origin using Taylor's infinite series, we can collect terms according to the powers of x and thus obtain a Maclaurin expansion of ϵ.

It is evident that in many cases the error can be reduced considerably by a suitable choice of coördinates; for example the relation

$$y - x \tan \sqrt[4]{x^2 + y^2} = 0$$

[16] We have used $f_m^{\,r}$ in order to take particular derivatives of f outside of the integral sign. Obviously this can be accomplished by means of expressions of the form $f^{r}(\xi)$ where ξ lies somewhere in the interval bounded by the two outermost points, since θ is non-negative; in fact the symbols $f_m^{\,r}$ and $f^{r}(\xi)$ can be interchanged in deriving expressions limiting $|\epsilon|$. With due regard for signs, an expression for ϵ can be derived which contains but a finite number of terms.

Pierpont: Theory of Functions of Real Variables, vol. 1, p. 366.

in cartesian coördinates becomes

$$\rho = \theta^2$$

in polar coördinates. As another example we may note that the substitution $x = \sin u$ will lead directly to the periodic interpolation function $P(\sin u)$.

8. **Special Cases of Balanced Systems.**[7] We now apply the theory developed in section 7, and consider the case where $i = 0, 1, 3, 5, \cdots$ odd, and $-\{c_{-1}^i\} = \{c_1^i\} = 1$; hence

$$S = [\{K_{-1}^0\}f_{-1} + \{K_1^0\}f_1] + \theta[\{K_{-1}^1\}f_{-1}^1 + \{K_1^1\}f_1^1] \\ + \theta^3[\{K_{-1}^3\}f_{-1}^3 + \{K_1^3\}f_1^3] + \cdots, \quad (90)$$

from which it follows that

$$\{K_1^i\} = \{U_1^i\} + \{V_1^i\}$$
$$\{K_{-1}^i\} = \{U_1^i\} - \{V_1^i\}.$$

The following systems of equations are due to (71) and (74):

$$\begin{aligned} \tfrac{1}{2} &= \{U_1^0\}\{0, -1\} + \{U_1^1\}\{0, 0\} + \{U_1^3\}\{0, 2\} + \cdots \\ \tfrac{1}{2}x^2 &= \{U_1^0\}\{2, -1\} + \{U_1^1\}\{2, 0\} + \{U_1^3\}\{2, 2\} + \cdots \\ \tfrac{1}{2}x^4 &= \{U_1^0\}\{4, -1\} + \{U_1^1\}\{4, 0\} + \{U_1^3\}\{4, 2\} + \cdots \\ &\cdots\cdots\cdots\cdots\cdots, \end{aligned} \quad (91)$$

$$\begin{aligned} \tfrac{1}{2}x &= \{V_1^0\}\{1, -1\} + \{V_1^1\}\{1, 0\} + \{V_1^3\}\{1, 2\} + \cdots \\ \tfrac{1}{2}x^3 &= \{V_1^0\}\{3, -1\} + \{V_1^1\}\{3, 0\} + \{V_1^3\}\{3, 2\} + \cdots \\ \tfrac{1}{2}x^5 &= \{V_1^0\}\{5, -1\} + \{V_1^1\}\{5, 0\} + \{V_1^3\}\{5, 2\} + \cdots \\ &\cdots\cdots\cdots\cdots\cdots. \end{aligned} \quad (92)$$

Noting that $\{n, m\} = 0$ if $(n - 1) < m$, we see that all terms of (91) that lie above the principal diagonal are zero; hence (91) is determinate, since its determinant for the first k equations is the product of the terms of the principal diagonal, and hence not zero. Noting that $\{n, m\} = \overline{m+1}!\binom{n}{m+1}$, we can factor $\overline{m+1}!$ out of the various columns of the determinant of the

first k equations of (92), which places $\Delta_{\epsilon,k}$ in the following form:

$$\Delta_{\epsilon,k}=(-1)^{k-1}\cdot\begin{vmatrix} 1 & 0 & \cdots & \cdot & 0 & 1 \\ \binom{3}{1} & 1 & \cdots & & 0 & 1 \\ \binom{5}{1} & \binom{5}{3} & \cdots & & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \binom{2k-1}{1} & \binom{2k-1}{3} & \cdots & \binom{2k-1}{2k-3} & 1 \end{vmatrix}$$

$$\times \prod_{i=1}^{i=k-1} \overline{2i-1}! = B_{2k}\frac{2^{2k}(2^{2k}-1)}{2k}\prod_{i=1}^{i=k-1}\overline{2i-1}! \neq 0;$$

hence (92) is determinate, since the Bernoullian numbers B_{2k} are known to differ from zero.[17] We thus know that the K's are determinate, together with $G(x)$; furthermore $\lambda = 2k$, there being k elements in S; and $|\epsilon|$ is limited by an expression of the form (89).

We shall now derive a special formula in which S contains only mid-terms; thus

$$S = \{K_0^0\}f_0 + \theta\{K_0^1\}f_0^1 + \theta^2\{K_0^2\}f_0^2 + \cdots + \theta^k\{K_0^k\}f_0^k.$$

The system for determining the K's is

$$\begin{aligned} 1 &= \{K_0^0\} \\ x &= 1!\{K_0^1\} \\ x^2 &= 2!\{K_0^2\} \\ &\cdot\ \ \cdot\ \ \cdot\ \ \cdot \\ x^k &= k!\{K_0^k\}; \end{aligned}$$

hence $\{K_0^i\} = \frac{x^i}{i!}$, $i = 0, 1, 2, \cdots k$. S contains $(k+1)$ single term elements; hence $\lambda = (k+1)$, and since $x^{k+1} = \sigma_{k+1} = G(x)$, we have finally [14, 16]

$$f(x) = f(0) + xf^1(0) + \frac{x^2f^2(0)}{2!} + \cdots + \frac{x^kf^k(0)}{k!} + \frac{x^{k+1}f^{k+1}(\xi)}{\overline{k+1}!}, \quad (93)$$

[17] *Pascal: Die Determinanten*, p. 136.
Pierpont: Functions of a Complex Variable, p. 289.

which is Maclaurin's expansion with the remainder. The substitution $F(x_1 + x) = f(x)$ places (93) in the usual form of Taylor's expansion.

9. **Possibility of Leaving c's as Unknown Constants.** In section 7 it was assumed that the c's were chosen arbitrarily and then fixed. This was done merely to fix ideas, for it is obvious that the theory developed is independent of the type of unknowns. We thus see that certain K's and c's may be fixed, subject to condition (62), and the remaining ones determined by placing derivatives of $\phi(\theta)$ equal to zero in the usual manner, the method being that of sections 2, 3, and 4. Where λ is the total number of unknowns, $|\epsilon| \leq \frac{f_m{}^i \sigma_i}{i!} + \cdots + \frac{f_m^{i+k} \omega_i}{i+k!}$ where $i = 1, 2, 3, \cdots \lambda$.

As an example let $S = \{K_0{}^0\} f_0 + \theta[\{K_{-1}{}^1\} f_{-1}{}^1 + \{K_1{}^1\} f_1{}^1]$ where $\{K_0{}^0\}$, $\{K_{-1}{}^1\}$, $\{K_1{}^1\}$, $\{c_{-1}{}^1\} = -\{c_1{}^1\}$ are to be determined. We have immediately

$$f(x) = f(0) + \frac{x}{4}\left[(2 - \sqrt{3}) f'\left(-\frac{x}{\sqrt{3}}\right) + (2 + \sqrt{3}) f'\left(\frac{x}{\sqrt{3}}\right)\right] + \epsilon$$

where

$$|\epsilon| \leq \frac{f_m{}^i \sigma_i}{i!} + \frac{f_m^{i+1} \tau_i}{i+1!} \quad i = 1, 2, 3, 4.$$

10. **Determinateness of Results.—Arbitrary Systems.** The theory developed in sections 7, 8, and 9 was developed entirely upon the basis of practical convenience, since the formulas that are of practical importance are special cases of it.

There is obviously no difficulty encountered in the general case where the number and choice of the elements of S are arbitrary, the number and choice of the terms of the various elements are without restriction, certain K's and c's are fixed arbitrarily, and those remaining are determined by placing successive derivatives of $\phi(\theta)$ equal to zero. Since in the derivation of (89) no use was made of (62), we can operate directly with $\phi^i(\theta)$ in the same manner as with (49), and thus obtain an expression of the form (89). The one restriction upon S is that the system which results for determining the unknowns is determinate. The case where there is but one

element in S was fully considered in section 2; however, the general case is much more difficult.

We shall consider here only the cases where the c's have been fixed and only K's are to be determined, other cases being left to trial.

Let us suppose that though the number of elements is arbitrary, no element is omitted; thus

$$i = 0, 1, 2, 3, \cdots I.$$

The system is determinate if and only if

$$\Delta \neq 0$$

where Δ is the determinant of the system, which after eliminating mid-terms by developing by minors is of the following form:

$$\Delta = B \begin{vmatrix} a_{10} & \cdots a_{10} & 0 & \cdots 0 & 0 & \cdots 0 & \cdots \\ a_{11}c_1 & \cdots a_{11}c_2 & a_{20} & \cdots a_{20} & 0 & \cdots 0 & \cdots \\ a_{12}c_1^2 \cdots & a_{12}c_2^2 & a_{21}c_3 & \cdots a_{21}c_4 & a_{30} & \cdots a_{30} & \cdots \\ a_{13}c_1^3 \cdots & a_{13}c_2^3 & a_{22}c_3^2 & \cdots a_{22}c_4^2 & a_{31}c_5 & \cdots a_{31}c_6 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1c_1^q \cdots & a_1c_2^q & a_2c_3^{q-1} \cdots & a_2c_4^{q-1} & a_3c_5^{q-2} \cdots & a_3c_6^{q-2} & \cdots \end{vmatrix} \tag{94}$$

where the a's and c's are constants representing the $\{r, i-1\}$'s and $\{c_j^i\}$'s of (71) respectively, and B is not zero. It is seen that each element contributes a block of terms to this determinant. We now develop this determinant by minors with respect to the first column, and so obtain a polynomial $P_1(c_1)$ that is of degree d_1 where $d_1 \leq q$. The c's of the first block are obviously roots of $P_1(c_1)$; however, as opposed to the case of (34), not all of the roots are known. As it is known that they exist, we write, using the factor theorem,

$$\frac{\Delta}{B} = P_1(c_1) = (c_1 - \alpha_{11})(c_1 - \alpha_{12}) \cdots (c_1 - \alpha_{1,\,d_1})\Delta_1 = \Pi_1\Delta_1 \tag{95}$$

where Δ_1 is the cofactor of $a_{1,\,d_1}c_1^{d_1}$, and it is seen that Δ_1 and the α_1's do not contain c_1. Since Δ_1 is a determinant of the same type as Δ, we may write

$$\Delta_1 = P_2(c_2) = (c_2 - \alpha_{21})(c_3 - \alpha_{22}) \cdots (c_2 - \alpha_{2,\,d_2})\Delta_2 = \Pi_2\Delta_2,$$

the procedure being that of (95), where Δ_2 and the α_2's contain neither c_1 nor c_2. Proceeding in this way we finally have

$$\Delta = B \cdot \Pi_1 \cdot \Pi_2 \cdot \Pi_3 \cdots \Pi_\omega \cdot V, \tag{96}$$

where development with respect to the upper row is made when the last column of each block is encountered; hence V is a Vandermonde determinant, which is not zero since the c's of each block are necessarily distinct. The α's of Π_ω are determined by the c's of V; hence once these c's are fixed we have a *discrete set of values which if assigned to the c of* Π_ω *will result in the vanishing of* Δ. Moreover, only these values of c will cause Δ to vanish. We now fix this c and immediately have a discrete set of values which cannot be assigned to the c of $\Pi_{\omega-1}$, and so on. We thus see that Δ *can be, but generally is not zero*; also that if Δ is zero in a particular case, an adjustment of the c's is all that is required to render the system determinate. It is therefore evident that S *need not be altered in form.*

Let us now suppose that k successive elements are omitted, there being elements on both sides of the gap so formed; then the block on the right has $(k + 1)$ more rows of zeros than has the adjacent block on the left. In deriving (96), successive columns of (94) were taken out, the degree of the determinant being lowered; however, we now continue this process only to the point where there are $(k + 1)$ columns to the left of the gap. Laplace's development then throws the determinant into the product of two others of the type (94). If there are several gaps, the process is continued only to the point where it is first possible to use Laplace's development, the resulting pair of determinants being considered separately as before. Thus we finally have either a product of gapless determinants each of the type (94), or the conclusion that Δ is zero.

A system is usually determinate if the first p equations involve no fewer than p unknowns. As an example, every S must have an element $\{E^0\}$.

As a further example, there can be no formula of the form

$$f(x) = [\{K_{-1}^0\} f(-c) + \{K_1^0\} f(c)] + \{K_0^1\} f'(0) + \epsilon,$$

for

$$\Delta = \begin{vmatrix} 1 & 1 & 0 \\ -c & c & 1 \\ c^2 & c^2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ c^2 & c^2 \end{vmatrix} = 0.$$

11. **Uniqueness of Results.** Let us suppose that by the above method there has been obtained an expression P, which is a special case of (41); and that the system which determines the unknowns of P consists of t equations, and is uniquely determinate. Also let us suppose that in some other way there has been derived a formula Π which gives exact results when applied to any polynomial of degree $\leq \delta$, where δ is a given number; and that Π is similar in form to P, by which we mean that the same unknowns were determined in both cases. It will now be shown that if

$$\delta \geq (t - 1), \tag{97}$$

then

$$P \equiv \Pi.$$

To this end we first replace every term of the form $Kf^i(c)$ in P and in Π by $K\theta^i f^i(c\theta)$, and thus obtain two corresponding expressions which we denote by S and $\sum$ respectively; then we place

$$\phi(\theta) = f(x\theta) - \sum,$$

and differentiate successively. Applying the law of the mean,

$$f^r(u\theta) = f^r(0) + u\theta f^{r+1}(\xi),$$

to refer each principal term of $\phi^r(\theta)$ to the origin, we have

$$\phi^r(\theta) = Q_r f^r(0) + \left\{ \begin{matrix} \text{Terms involving higher} \\ \text{derivatives of } f. \end{matrix} \right\} \tag{98}$$

where Q_r involves only the unknowns, which were determined by the other method. We now integrate (98) successively between the limits 0 and θ, using the law of the mean [16] upon the higher order terms to keep mean values of the derivatives of f outside of the integral sign. The result is

$$\phi(\theta) = \frac{Q_r \theta^r f^r(0)}{r!} + \Omega_r \tag{99}$$

where Ω_r represents terms involving higher derivatives of f. By

hypothesis we know that if

$$f(x) = P_r(x) = \textit{Any polynomial of degree } r \tag{100}$$

with the condition that $r \leq \delta$, we have from (99) that

$$\textit{Error} = \phi(1) = 0 = \frac{Q_r P_r{}^r(0)}{r!}, \tag{101}$$

since Ω_r vanishes due to the order of the derivatives which it involves. Applying (101), we choose $P_r(x)$ so that $P_r{}^r(0) \neq 0$; hence (101) becomes

$$Q_r = 0 \qquad \text{if} \qquad r \leq \delta.$$

Placing $r = 0, 1, 2, 3, \cdots \delta$, we obtain a system of $(\delta + 1)$ equations that are satisfied by the unknowns of Π. But the unknowns of P were *determined* by the relations

$$Q_r = 0 \qquad \text{where} \qquad r = 0, 1, 2, 3, \cdots (t - 1);$$

therefore

$$P \equiv \Pi \tag{102}$$

if and only if

$$\delta \geq (t - 1).$$

As an example, $t = n$ in the case of (35), whereas $\delta = (n - 1)$ in the case of the corresponding formula of Lagrange; hence the corresponding formulas are identical, and the title of section 2 is justified.

12. **Case Where Points of the Same Element Approach Coincidence.** To fix ideas let us consider (33) in the case where $c_j \to c_i$. Since the function

$$\epsilon(x) = f(x) - P(x)$$

vanishes at the points c_i and c_j, we know by Rolle's Theorem that $\epsilon^1(x)$ vanishes at some point ξ, where $c_i < \xi < c_j$; however, when $c_j \to c_i$, $\xi \to c_i$; hence we have finally

$$f^1(c_i) = P^1(c_i),$$

and $P(x)$ must now coincide with the polynomial that is derived

from the expression

$$S = \sum_{\substack{u=1 \\ u \neq j}}^{u=n} \{K_u^0\} f_u + \theta \{K_i^1\} f_i',$$

since the condition (102) is satisfied. This fact can also be shown directly by expressing the K's of (33) as the quotient of two determinants. c_j may be taken as an independent variable and L'Hospital's rule applied, since $K_k \to 0/0$, $k \neq i, j$; thus the elements of the jth column are replaced by their derivatives with respect to c_j at the point c_i, and $K_k = \{K_k^0\}$ identically when $k \neq i, j$. When $k = i$ and j, we collect the two resulting terms of P; thus

$$K_i f(c_i) + K_j f(c_j)$$

$$= \frac{f(c_i)\begin{vmatrix} \cdots 1 & 1 & \cdots \\ \cdots x & c_i + e & \cdots \\ \cdots x^2 & c_i^2 + 2\xi_1 e & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} - [f(c_i) + f^1(\xi)e] \begin{vmatrix} \cdots 1 & 1 & \cdots \\ \cdots x & c_i & \cdots \\ \cdots x^2 & c_i^2 & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix}}{\begin{vmatrix} \cdots 1 & 0 & \cdots \\ \cdots c_i & 1 & \cdots \\ \cdots c_i^2 & 2\xi_1 & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} e}$$

where $e = (c_j - c_i)$ and $c_i < \xi_1, \xi < c_j$; hence

$$K_i f(c_i) + K_j f(c_j) \to \{K_i^0\} f(c_i) + \{K_i^1\} f^1(c_i),$$

and the two polynomials are identical.

There is evidently no difficulty to be encountered in deriving analogous relations when groups of points of one or more elements approach common limits. In particular, when all of the points of $P(x)$ approach a common limit, $P(x)$ becomes a Taylor's expansion of $f(x)$ about that limit.

13. **Relation Between Formulas for the Approximation of Integrals, and the Corresponding Interpolation Polynomials.** Let us consider the interpolation polynomial $P(x)$ in the derivation of which λ unknowns were determined; then in general

$$f(x) = P(x) + \epsilon \qquad \text{where} \qquad |\epsilon| \leq \frac{f_m^{\lambda} \sigma_{\lambda}}{\lambda!} + \cdots + \frac{f_m^{\lambda+k} \omega_{\lambda}}{\overline{\lambda + k}!}.$$

Integrating, we have

$$\int_{-\theta}^{\theta} f(x)dx = \int_{-\theta}^{\theta} P(x)dx + \int_{-\theta}^{\theta} \epsilon dx;$$

hence if

$$f(x) = P_\alpha(x) = \textit{Any polynomial of degree } \alpha,$$

$$\int_{-\theta}^{\theta} \epsilon dx = 0 \quad \text{if} \quad \alpha \leqq \delta = (\lambda - 1).$$

It is thus seen that $\int_{-\theta}^{\theta} P(x)dx$ gives an exact result when used to approximate $\int_{-\theta}^{\theta} P_\alpha(x)dx$ where $\alpha \leqq \delta$.

In the derivation of the corresponding integration formula, λ equations were involved; hence placing $\lambda = t$, we have

$$\delta = (t - 1),$$

which proves that *the integral of an interpolation polynomial is identical with the corresponding formula of approximate integration, the limits being the same in both cases.*[17, 18]

As an example, Cotes' coefficients for $(2n + 1)$ points are obtained by integrating (22), the integral being of a familiar form. Also, we have from (23)

$$\begin{aligned} [K_i] &= \int_{-nh}^{nh} K_i dx = \int_{-nh}^{nh} U_i dx + \int_{-nh}^{nh} V_i dx = [U_i] + [V_i], \\ [K_{-i}] &= \int_{-nh}^{nh} K_{-i} dx = \int_{-nh}^{nh} U_i dx - \int_{-nh}^{nh} V_i dx = [U_i] - [V_i], \end{aligned} \tag{103}$$

which placed in the integral of (24) gives

$$\begin{aligned} \sum_{j=1}^{j=n} [U_i] j^i &= \int_{-nh}^{nh} \frac{1}{2} \left(\frac{x}{h}\right)^i dx = \frac{hn^{i+1}}{(i+1)} \quad i = 2, 4, 6, \cdots 2n, \\ \sum_{j=1}^{j=n} [V_i] j^i &= \int_{-nh}^{nh} \frac{1}{2} \left(\frac{x}{h}\right)^i dx = 0 \quad i = 1, 3, 5, \cdots (2n-1). \end{aligned} \tag{104}$$

Since the determinants of systems (104) are the same as those of

[17] *Crout: A Method for Deriving Formulas for the Approximate Calculation of Integrals*, this *Journal*, vol. 7, p. 159.

(24), they are different from zero; hence

$$[V_j] = 0 \qquad j = 1, 2, 3, \cdots n,$$

and

$$[K_{-j}] = [U_j] = [K_j] \qquad j = 1, 2, 3, \cdots n, \tag{105}$$

due to (103). The first equation of (24) becomes

$$[K_0] = 2\left\{nh - \sum_{j=1}^{j=n} [K_j]\right\},$$

and Cotes' coefficients are thus completely determined.[7, 19]

[19] *Rutledge: The Explicit Determination of Cotes' Coefficients for Polynomial Area*, this *Journal*, vol. 1, p. 78. It should be noted that the $[K]$'s of (105) are $2nh$ times the K's of this reference.

DECOMPOSITION OF DETERMINANTS

By Lepine Hall Rice

In a former paper [1] we dealt with the process of decomposition —the expression of a determinant of higher class as an algebraic sum of determinants of any desired lower class. We generalized the kind of decomposition previously known, which we shall call *straight* decomposition, and also gave a new form which we called *crossed* decomposition. We shall now improve the treatment of straight decomposition and show that by repeating this process all the results of crossed decomposition formerly given are obtained more simply than before and that additional results of this kind are obtained. The inclusive formula here reached is in consequence both broader and more easily comprehended and applied than are the previous formulas.

§1. Straight Decomposition

To make a straight decomposition, we took a p-way determinant of any signancy and of order n, and considered its k-way couches [2] of any given aspect. We assembled n perjunctive couches as the layers (of the first direction) of a $(k+1)$-way matrix of order n, and formed all the $(n!)^{p-k-1}$ possible matrices of this description. We took the determinant, of a certain signancy, or the permanent, as the case might be, of each of these matrices, and prefixed a specified sign to each determinant or permanent. The sum of these *components* was shown to be equal to the p-way determinant.

We now denote the nonsignant and signant indices in the given aspect by α's and β's respectively, and the remaining

[1] P-way determinants, *Am. J. of Math.*, vol. XL (1918), p. 242, at p. 247 and p. 252. See also M. Lecat, Coup d'oeil sur la théorie des déterminants supérieurs dans son état actuel, Lamertin (Bruxelles), 1927, p. 49.

[2] A couche (called in the former paper a sublayer) of aspect $i_1 i_2 \cdots i_{p-k}$ consists of those elements of the matrix in which the indices $i_1, i_2, \cdots, i_{p-k}$ have each a fixed value; n couches of this aspect are *perjunctive* if no two values of any index are the same, so that the values of each index form a permutation of $12 \cdots n$.

nonsignant and signant indices by γ's and δ's respectively, and write [3] the p-way matrix

$$\| A_{\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_s \gamma_1 \cdots \gamma_t \delta_1 \cdots \delta_u} \|, \quad \begin{matrix} r+s=p-k, \\ t+u=k, \end{matrix} \tag{1}$$

and the determinant

$$|A| = |A_{\overset{+}{\alpha_1} \cdots \overset{\pm}{\alpha_r} \beta_1 \cdots \beta_s \overset{+}{\gamma_1} \cdots \overset{\pm}{\gamma_t} \delta_1 \cdots \delta_u}|. \tag{2}$$

Denote by

$$\| A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_s)\gamma_1 \cdots \gamma_t \delta_1 \cdots \delta_u} \| \tag{3}$$

any one of the $(k+1)$-way matrices, $\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_s$ being a $(p-k)$-partite index that takes n perjunctive values.

If s is odd, make the γ's in (3) nonsignant, the δ's signant, and the $(p-k)$-partite index *signant on* β_s, that is, give the $(p-k)$-partite range the sign of the β_s-range; [4] prefix to each of the $(n!)^{p-k-1}$ determinants thus obtained the product of the signs of the remaining β-ranges (if any) in the transversals of that determinant (this product being constant in all the transversals of that determinant, and in each transversal for all arrangements of its elements, since $s-1$ is even); and take the sum of these components.

If s is even, make the γ's nonsignant, the δ's signant, and the $(p-k)$-partite index nonsignant; prefix to each determinant (or permanent if there are no δ's) the product of the signs of the β-ranges (if any) in its transversals; and take the sum of these components.

In each case every transversal thus finally receives the same sign that it has in $|A|$.

The decomposition may be written, [5] if s is odd,

$$|A| = \sum_{\alpha\beta}^{(n!)^{p-k-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_{s-1}} |A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_{s-1}\overset{\pm}{\beta_s})\overset{+}{\gamma_1} \cdots \gamma_t \overset{\pm}{\delta_1} \cdots \delta_u}|; \tag{4}$$

[3] Indices will be arranged and rearranged as convenient, when generality is not thereby lessened.

[4] This may be any one of the β-ranges.

[5] In the former paper it was specified that the sign-factors here denoted by ϵ_{β_1}, etc. (there by $\pm_{\beta_1}$, etc.) should be fixed *when the* β_s*-range read* $12 \cdots n$. But as pointed out above, the *product* $\epsilon_{\beta_1} \cdots \epsilon_{\beta_{s-1}}$ (in (4)) or $\epsilon_{\beta_1} \cdots \epsilon_{\beta_s}$ (in (5)) is independent of the order of values in the β_s-range. This frees us from the necessity of thinking of any definite order of the latter. That is sometimes an advantage, precisely as it is an advantage to define the signs of the terms of a 2-way determinant without requiring the elements of the term to be in any special order.

and if s is even $(0, 2, \cdots)$,

$$|A| = \sum_{\alpha\beta}^{(n!)^{p-k-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_s} |A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_s)} \overset{+}{\gamma_1 \cdots \gamma_t} \overset{\pm}{\delta_1 \cdots \delta_u}|. \quad (5)$$

If there are no β's and no δ's, we have the decomposition of a permanent.

Complete decomposition into 2-way components occurs when the couches are files. In (4) there will then be one δ and no γ's, and files of a signant direction will be formed into determinants:

$$|A| = \sum_{\alpha\beta}^{(n!)^{p-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_{p-1}} |A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_{p-1}) \delta}^{\pm \pm}|; \quad (6)$$

while in (5) there will be one γ and no δ's, and files of a nonsignant direction will be formed into permanents:

$$|A| = \sum_{\alpha\beta}^{(n!)^{p-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_s} |A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \beta_s)_\gamma}^{++}|. \quad (7)$$

The decompositions (necessarily complete) of 3-way determinants and permanents are:

$$|A_{\alpha\beta\delta}^{+\pm\pm}| = \sum_{\alpha\beta}^{n!} |A_{(\alpha\beta)\delta}^{\pm\pm}|, \quad (8)$$

$$|A_{\beta_1\beta_2\gamma}^{\pm\pm+}| = \sum_{\beta}^{n!} \epsilon_{\beta_1} \epsilon_{\beta_2} |A_{(\beta_1\beta_2)\gamma}^{++}|, \quad (9)$$

$$|A_{\alpha_1\alpha_2\gamma}^{+++}| = \sum_{\alpha}^{n!} |A_{(\alpha_1\alpha_2)\gamma}^{++}|. \quad (10)$$

§2. Crossed Decomposition and the General Formula

Let us apply formula (4) to itself. Selecting a couche-aspect

$$\alpha'_1 \cdots \alpha'_r \beta'_1 \cdots \beta'_{s'}, \quad (11)$$

the α''s and β''s being chosen from the γ's and δ's, and the remaining γ's and δ's being denoted by

$$\gamma'_1 \cdots \gamma'_{t'} \delta'_1 \cdots \delta'_{u'}, \quad (12)$$

we obtain the crossed decomposition

$$|A| = \sum_{\alpha\beta,\,\alpha'\beta'}^{(n!)^{r+s-1+r'+s'-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_{s-1}} \epsilon_{\beta'_1} \cdots \epsilon_{\beta'_{s'-1}}$$

$$\cdot \left| A_{(\alpha_1 \cdots \alpha_r \overset{\pm}{\beta_1} \cdots \beta_{s-2}\beta_s)(\alpha'_1 \cdots \alpha'_{r'} \overset{\pm}{\beta'_1} \cdots \beta'_{s'-2}\beta'_{s'})} \overset{+}{\gamma'_1 \cdots \gamma'_{t'}} \overset{\pm}{\delta'_1 \cdots \delta'_{u'}} \right|. \quad (13)$$

Similarly applying formula (5) to itself we have

$$|A| = \sum_{\alpha\beta,\,\alpha'\beta'}^{(n!)^{r+s-1+r'+s'-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_s} \epsilon_{\beta'_1} \cdots \epsilon_{\beta'_{s'}}$$

$$\cdot \left| A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \overset{+}{\beta_s})(\alpha'_1 \cdots \alpha'_{r'} \beta'_1 \cdots \overset{+}{\beta'_{s'}})} \overset{+}{\gamma'_1 \cdots \gamma'_{t'}} \overset{\pm}{\delta'_1 \cdots \delta'_{u'}} \right|. \quad (14)$$

And again, applying (5) to (4) we have

$$|A| = \sum_{\alpha\beta,\,\alpha'\beta'}^{(n!)^{r+s-1+r'+s'-1}} \epsilon_{\beta_1} \cdots \epsilon_{\beta_{s-1}} \epsilon_{\beta'_1} \cdots \epsilon_{\beta'_{s'}}$$

$$\cdot \left| A_{(\alpha_1 \cdots \alpha_r \beta_1 \cdots \overset{\pm}{\beta_{s-1}\beta_s})(\alpha'_1 \cdots \alpha'_{r'} \beta'_1 \cdots \overset{+}{\beta'_{s'}})} \overset{+}{\gamma'_1 \cdots \gamma'_{t'}} \overset{\pm}{\delta'_1 \cdots \delta'_{u'}} \right|. \quad (15)$$

In (13) and (14) if there are no γ''s and no δ''s we have complete decompositions, of which (6) and (7) are special cases.

These new components may in turn be decomposed, and so on. The p indices are finally partitioned into q multipartite indices, some (say g) of which may contain an odd number of originally signant indices while the rest contain an even number. If the jth multipartite index be written

$$\alpha_{j1} \cdots \alpha_{jr_j} \beta_{j1} \cdots \beta_{js_j}, \quad \text{or} \quad \begin{matrix}[\alpha][\beta], \\ j1 \quad j1 \\ \cdots \quad \cdots \\ jr_j \quad js_j\end{matrix} \quad (16)$$

and $s_1, \cdots, s_g$ be odd and $s_h, \cdots, s_q (h = g + 1)$ be even, and the determinant be written

$$|A| = \left| A \begin{matrix} \overset{+}{[\alpha]} & \overset{\pm}{[\beta]} & \cdots & \overset{+}{[\alpha]} & \overset{\pm}{[\beta]} & \overset{+}{[\alpha]} & \overset{\pm}{[\beta]} & \cdots & \overset{+}{[\alpha]} & \overset{\pm}{[\beta]} \\ 11 & 11 & & g1 & g1 & h1 & h1 & & q1 & q1 \\ \cdots & \cdots & & \cdots & \cdots & \cdots & \cdots & & \cdots & \cdots \\ 1r_1 & 1s_1 & & gr_g & gs_g & hr_h & hs_h & & qr_q & qs_q \end{matrix} \right|, \quad (17)$$

then we have as the general formula for decomposition into

q-way components

$$|A| = \sum_{\alpha\beta}^{(n!)} \left(\prod_{[\beta]}^{\Sigma r + \Sigma s - q} \epsilon_\beta \right) |A \begin{pmatrix} [\alpha] & [\beta] & \overset{\pm}{\beta} \\ 11 & 11 & 1s_1 \\ \cdots & \cdots & \\ 1r_1 & 1, s_1 - 1 & \end{pmatrix} \cdots \begin{pmatrix} [\alpha] & [\beta] & \overset{\pm}{\beta} \\ q1 & q1 & qs_q \\ \cdots & \cdots & \\ qr_q & q, s_q - 1 & \end{pmatrix} \begin{pmatrix} [\alpha] & [\beta] \\ h1 & h1 \\ \cdots & \cdots \\ hr_h & hs_h \end{pmatrix}^{+} \cdots \begin{pmatrix} [\alpha] & [\beta] \\ q1 & q1 \\ \cdots & \cdots \\ qr_q & qs_q \end{pmatrix}^{+} |, \tag{18}$$

where all the β-ranges in brackets contribute to the sign of each component.

It is evident that all the transversals of all the components are all the transversals of $|A|$ and have the signs that they have in $|A|$.

§3. A List of Decompositions

The following list gives the decompositions of 4-way, 5-way, and 6-way determinants and permanents. For class p the main grouping is $q = p - 1$, $q = p - 2$, etc., with subgroups based on the possible partitions of the p indices, where more than one partition will give the same value of q. The decompositions for class 4 are presented in the notation used above for class 3, and also in an abridged notation in which the indices are denoted simply by their signancy marks, the signancy of the components is put above, the ϵ's to be prefixed are indicated by placing them over their indices, and the summation sign is omitted. The decompositions for classes 5 and 6 are given only in this abridged notation.

Of 4-way into 3-way, with the partition 2–1–1:

$$|A^{\pm\pm\pm\pm}_{\beta_1\beta_2\delta_1\delta_2}| = \sum_{\beta}^{n!} \epsilon_{\beta_1}\epsilon_{\beta_2} |A_{(\beta_1\beta_2)}{}^{+\pm\pm}_{\delta_1\delta_2}|, \quad \text{or} \quad {}^{\cdot\ \cdot}_{(\pm\pm)}{}^{+\pm\pm}_{\pm\pm} \tag{19}$$

$$|A^{\pm\pm++}_{\beta_1\beta_2\gamma_1\gamma_2}| = \sum_{\beta}^{n!} \epsilon_{\beta_1}\epsilon_{\beta_2} |A_{(\beta_1\beta_2)}{}^{+++}_{\gamma_1\gamma_2}|, \quad \text{or} \quad {}^{\cdot\ \cdot}_{(\pm\pm)}{}^{+++}_{++} \tag{20}$$

$$|A^{+\pm+\pm}_{\alpha\beta\gamma\delta}| = \sum_{\alpha\beta}^{n!} |A_{(\alpha\beta)}{}^{\pm}_{\gamma}{}^{+\pm}_{\delta}|, \quad \text{or} \quad {}^{\pm}_{(+\pm)}{}^{+\pm}_{+\pm} \tag{21}$$

$$|A^{++\pm\pm}_{\alpha_1\alpha_2\delta_1\delta_2}| = \sum_{\alpha}^{n!} |A_{(\alpha_1\alpha_2)}{}^{+\pm\pm}_{\delta_1\delta_2}|, \quad \text{or} \quad {}^{+\pm\pm}_{(++)}{}^{+\pm\pm}_{\pm\pm} \tag{22}$$

$$|A^{+++++}_{\alpha_1\alpha_2\gamma_1\gamma_2}| = \sum_{\alpha}^{n!} |A_{(\alpha_1\alpha_2)}{}^{+}_{\gamma_1}{}^{++}_{\gamma_2}|, \qquad \text{or} \quad {}^{\cdot\,\cdot}_{(++)}{}^{+++}_{++} \tag{23}$$

Of 4-way into 2-way, with the partition 3–1:

$$|A^{\pm\pm\pm\pm}_{\beta_1\beta_2\beta_3\delta}| = \sum_{\beta}^{(n!)^3} \epsilon_{\beta_1}\epsilon_{\beta_2} |A_{(\beta_3\beta_1\beta_2)}{}^{\pm}_{\delta}|, \qquad \text{or} \quad {}^{\cdot\,\cdot}_{(\pm\pm\pm)}{}^{\pm}_{\pm} \tag{24}$$

$$|A^{+\pm\pm+}_{\alpha\beta_1\beta_2\gamma}| = \sum_{\alpha\beta}^{(n!)^3} \epsilon_{\beta_1}\epsilon_{\beta_2} |A_{(\alpha\beta_1\beta_2)}{}^{++}_{\gamma}|, \qquad \text{or} \quad {}^{\cdot\,\cdot}_{(+\pm\pm)}{}^{++}_{+} \tag{25}$$

$$|A^{++\pm\pm}_{\alpha_1\alpha_2\beta\delta}| = \sum_{\alpha\beta}^{(n!)^3} |A_{(\alpha_1\alpha_2\beta)}{}^{\pm}_{\delta}|, \qquad \text{or} \quad {}^{\pm}_{(++\pm)}{}^{\pm}_{\pm} \tag{26}$$

$$|A^{+++++}_{\alpha_1\alpha_2\alpha_3\gamma}| = \sum_{\alpha}^{(n!)^3} |A_{(\alpha_1\alpha_2\alpha_3)}{}^{++}_{\gamma}|, \qquad \text{or} \quad {}^{\cdot\,\cdot}_{(+++)}{}^{++}_{+} \tag{27}$$

Of 4-way into 2-way, with the partition 2–2:

$$|A^{\pm\pm\pm\pm\pm\pm}_{\beta_1\beta_2\beta'_1\beta'_2}| = \sum_{\beta,\beta'}^{n!n!} \epsilon_{\beta_1}\epsilon_{\beta_2}\epsilon_{\beta'_1}\epsilon_{\beta'_2} |A_{(\beta_1\beta_2)\,(\beta'_1\beta'_2)}{}^{+}_{+}|, \qquad \text{or} \quad {}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+}_{+}\;{}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+}_{+} \tag{28}$$

$$|A^{\pm\pm\pm+}_{\beta_1\beta_2\alpha'_1\alpha'_2}| = \sum_{\beta,\alpha'}^{n!n!} \epsilon_{\beta_1}\epsilon_{\beta_2} |A_{(\beta_1\beta_2)}{}^{+}_{(\alpha'_1\alpha'_2)}{}^{+}_{+}|, \qquad \text{or} \quad {}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+}_{+}\;{}^{\pm}_{(++)}{}^{+}_{+} \tag{29}$$

$$|A^{+\pm+\pm}_{\alpha\beta\alpha'\beta'}| = \sum_{\alpha\beta,\,\alpha'\beta'}^{n!n!} |A_{(\alpha\beta)\,(\alpha'\beta')}|, \qquad \text{or} \quad {}^{\pm}_{(+\pm)}\;{}^{\pm}_{(+\pm)} \tag{30}$$

$$|A^{++++}_{\alpha_1\alpha_2\alpha'_1\alpha'_2}| = \sum_{\alpha,\,\alpha'}^{n!n!} |A_{(\alpha_1\alpha_2)}{}^{+}_{(\alpha'_1\alpha'_2)}{}^{+}_{+}|, \qquad \text{or} \quad {}^{+}_{(++)}\;{}^{+}_{(++)} \tag{31}$$

Of 5-way into:

4-way:

$${}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+\pm\pm}_{+\pm\pm} \tag{32}$$

$${}^{\pm}_{(+\pm)}\;{}^{\pm\pm\pm}_{\pm\pm\pm} \tag{33}$$

$${}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+++}_{+++} \tag{34}$$

$${}^{\pm}_{(+\pm)}\;{}^{+\pm\pm}_{+\pm\pm} \tag{35}$$

$${}^{+}_{(++)}\;{}^{+\pm\pm}_{+\pm\pm} \tag{36}$$

$${}^{+}_{(++)}\;{}^{+++}_{+++} \tag{37}$$

3-way:

$${}^{\cdot\,\cdot}_{(\pm\pm\pm)}{}^{\pm}_{+\pm} \tag{38}$$

$${}^{\cdot\,\cdot}_{(+\pm\pm)}{}^{+\pm\pm}_{\pm\pm} \tag{39}$$

$${}^{\cdot\,\cdot}_{(+\pm\pm)}{}^{+++}_{++} \tag{40}$$

$${}^{\pm}_{(++\pm)}\;{}^{+\pm}_{+\pm} \tag{41}$$

$${}^{+}_{(+++)}\;{}^{+\pm\pm}_{\pm\pm} \tag{42}$$

$${}^{+}_{(+++)}\;{}^{+++}_{++} \tag{43}$$

$${}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+}_{+}\;{}^{\cdot\,\cdot}_{(\pm\pm)}{}^{++}_{+} \tag{44}$$

$${}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+}_{+}\;{}^{\pm}_{(+\pm)}\;{}^{\pm}_{\pm} \tag{45}$$

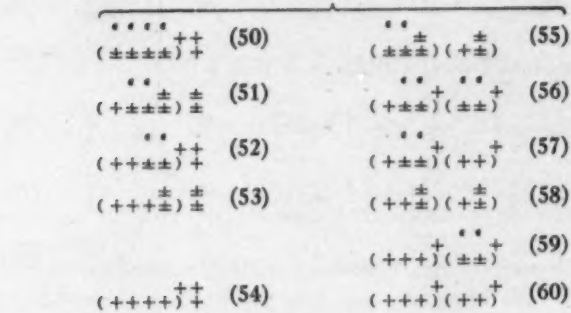
$${}^{\cdot\,\cdot}_{(\pm\pm)}{}^{+}_{+}\;{}^{+}_{(++)}\;{}^{++}_{+} \tag{46}$$

$${}^{\pm}_{(+\pm)}\;{}^{\pm}_{(+\pm)}\;{}^{+}_{+} \tag{47}$$

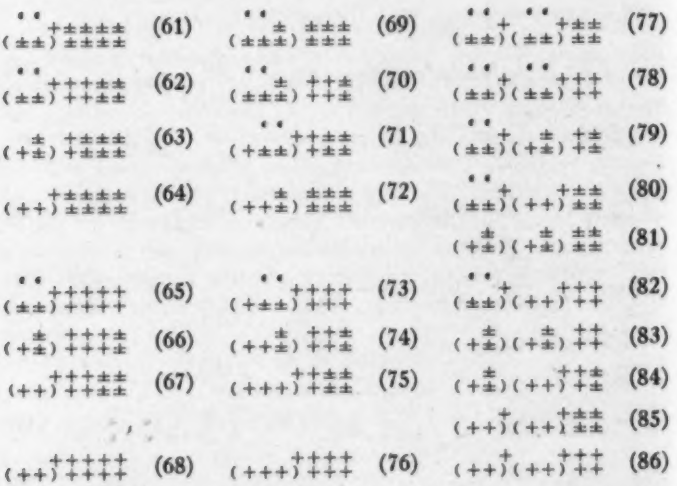
$${}^{\pm}_{(+\pm)}\;{}^{+}_{(++)}\;{}^{\pm\pm}_{\pm} \tag{48}$$

$${}^{+}_{(++)}\;{}^{+}_{(++)}\;{}^{++}_{+} \tag{49}$$

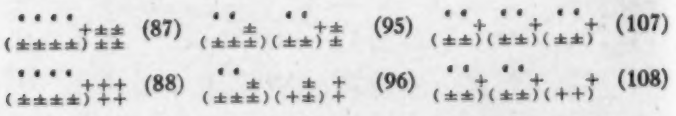
2-way:



Of 6-way into 5-way and into 4-way:



Of 6-way into 3-way:



 (89)
 (97)
(109)

 (90)
 (98)

 (99)

(100)

 (91)
(101)
 (110)

(92)
(102)
 (111)

(93)
(103)

 (104)

(105)

 (94)
 (106)
(112)

Of 6-way into 2-way:

 (113)
 (119)
(127)

 (114)
 (120)
 (128)

(115)
 (121)
 (129)

 (122)

 (116)
 (123)
(130)

(117)
(124)
(131)

(125)

 (118)
 (126)
(132)

§4. Conclusion

In all classes from class 4 up, some of the decompositions given by the general formula (18) were not obtainable by the formulas of the previous paper; of these, (28) is the simplest.

The equivalents of (29), (30), and (31) were formerly obtainable. This shows the incorrectness of a statement in the literature, that one cannot have crossed decomposition if the class is less than 5.

PRINCIPAL DIRECTIONS IN A SUBSPACE V_m IMMERSED IN A RIEMANNIAN SPACE V_n

By C. L. E. Moore

1. Ricci [1] showed that in a Riemannian space V_m the directions for which the mean curvature has an extreme value are the principal directions of a symmetric tensor (Ricci tensor) of the second order. These directions, m in number, are mutually perpendicular and do not depend on the containing space. It is the object of this note to show that there are two other sets of m mutually perpendicular directions passing through each point of V_m. These latter depend on the enveloping space just as the lines of curvature do for a hypersurface. Besides the Riemann tensor for V_m we have in V_m two other tensors of the fourth order satisfying the same identities as the Riemann tensor. One is the V_m-component of the Riemann tensor for V_n and the other is given by the sum of the second order minors of the second fundamental forms of V_m. Contracting these in the middle with the first fundamental tensor of V_m we obtain two symmetric tensors of the second order. From these we get our two sets of principal directions.

2. Let the Riemannian space V_n be defined by

$$ds^2 = a_{rs}dx^r dx^s, \qquad r, s = 1, 2, \cdots n \tag{1}$$

and the subspace V_m by

$$ds^2 = a_{rs}'dx^r dx^s, \qquad r, s = 1, 2, \cdots m. \tag{2}$$

If an orthogonal n-tuple λ_i be chosen so that $\lambda_1, \lambda_2, \cdots \lambda_m$ lie in V_m, then the coefficients of the two fundamental forms can be

[1] Direzioni e invariante principali in una varieta qualunque. Atti del reale Instituto Veneto, Vol. 63, pp. 1233–1239. See also, Struik, Grundzüge der mehrdimensionalen Differenzialgeometrie, page 67, Eisenhart, Riemannian Geometry, page 114.

expressed as follows,

$$a_{rs} = \sum_{1}^{n} {}_{i}\lambda_{i|r}\lambda_{i|s},$$

$$a'_{rs} = \sum_{1}^{m} {}_{i}\lambda_{i|r}\lambda_{i|s},$$

where

$$\sum {}_{i}\lambda_{i|r}\lambda_{i|}{}^{s} = 0, \qquad r \neq s$$

$$\sum {}_{i}\lambda_{i|r}\lambda_{i|}{}^{r} = 1.$$

The V_m-component of any tensor $X_{rs}{}^{t}$ lying in V_n can be obtained [2] by contracting with

$$\Delta^{rsh}_{ij\,l} = \sum_{1}^{m}{}_{\alpha}\lambda_{\alpha|}{}^{r}\lambda_{\alpha|\,i}\sum_{1}^{m}{}_{\beta}\lambda_{\beta|}{}^{s}\lambda_{\beta|\,j}\sum_{1}^{m}{}_{\gamma}\lambda_{\gamma|}{}^{t}\lambda_{\gamma|\,k}.$$

The V_m components of the Riemann tensor R_{mnpq} of V_n at a point of V_m are

$$R''_{\alpha\beta\gamma\delta} = \Delta^{mnpq}_{\alpha\beta\gamma\delta}R_{mnpq} \tag{3}$$

The Riemannian curvature of V_n for an orientation $\lambda_{h|}$, $\lambda_{k|}$ is

$$r_{hk} = R_{\alpha\beta\gamma\delta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta}\lambda_{h|}{}^{\gamma}\lambda_{k|}{}^{\delta}$$

and if the directions $\lambda_{h|}$, $\lambda_{k|}$ lie in the V_m this can evidently be written in the form,

$$r_{hk} = r''_{hk} = R''_{\alpha\beta\gamma\delta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta}\lambda_{h|}{}^{\gamma}\lambda_{k|}{}^{\delta}. \tag{4}$$

If this be summed on h from 1 to m we have

$$r''_{k} = R''_{\alpha\beta\gamma\delta}a'^{\alpha\gamma}\lambda_{k|}{}^{\beta}\lambda_{k|}{}^{\delta} = R''_{\beta\delta}\lambda_{k|}{}^{\beta}\lambda_{k|}{}^{\delta} \tag{5}$$

where

$$R''_{\beta\delta} = R''_{\alpha\beta\gamma\delta}a'^{\alpha\gamma}. \tag{6}$$

This last tensor is obviously symmetric and lies wholly in V_m. The quantity r''_{k} is seen to be independent $\lambda_{h|}$ and to depend on $\lambda_{k|}$ only. It is the sum of all the Riemann curvatures (with respect to (1)) of the $m-1$ orientations, lying in V_m, and determined by $\lambda_{k|}$ and one of the $m-1$ mutually perpendicular directions lying in V_m. We will call this invariant the *mean Riemannian curvature* with respect to V_n of V_m for the direction $\lambda_{k|}$.

[2] Struik, loc. cit., page 31.

Summing (5) on k from 1 to m

$$R'' = R''_{\alpha\beta\gamma\delta}a'^{\alpha\gamma}a'^{\beta\delta}. \tag{7}$$

This invariant is the sum of all curvatures (with respect to (1)) of the orientations determined by the orthogonal m-tuple in V_m. Since r''_k is a part of R'' this latter must break up into two invariants of which one is r''_k and the other is

$$r'''_k = \sum_1^m {}_{ij}\, r''_{ij} \qquad i, j \neq k.$$

Thus if $m = 4$, and $k = 1$,

$$r''_1 = r''_{12} + r''_{13} + r''_{14}$$
$$r'''_1 = r''_{23} + r''_{24} + r''_{34}.$$

These two invariants depend on the direction $\lambda_{1|}$. Since $R''_{\beta\delta}$ is symmetric its principal directions can be determined as the solutions of the equations

$$(R''_{\beta\delta} - r''_k a'_{\beta\delta})\lambda_{k|}{}^{\beta} = 0.$$

These principal directions are known to be mutually perpendicuar and the resulting values of r''_k are known to be the maximum or minimum values of the mean curvature with respect to V_n. Hence *The m directions through a point of V_m for which the mean curvature with respect to V_n has an extreme value are mutually perpendicular.*

These directions are in general different from the Ricci principal directions determined by the Riemann tensor $R'_{\alpha\beta\gamma\delta}$ of the space V_m.

3. Struik [3] expressed the Gauss relation, for a subspace V_m of V_n, in the following form

$$\sum_1^{n-m} {}_i\, h_{i|\alpha\gamma}h_{i|\beta\delta} - h_{i|\alpha\delta}h_{i|\beta\gamma} = R'_{\alpha\beta\gamma\delta} - R''_{\alpha\beta\gamma\delta} \tag{8}$$

where $h_{i|\alpha\gamma}\cdots$ are the coefficients of the second fundamental form of V_m, $R'_{\alpha\beta\gamma\delta}$ are the Riemann symbols formed from (2) and $R''_{\alpha\beta\gamma\delta}$ the V_m-components of those formed from (1). De-

[3] Loc. cit., page 124.

noting the left side of (8) by $H_{\alpha\beta\gamma\delta}$ and contracting both sides with $\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta}\lambda_{h|}{}^{\gamma}\lambda_{k|}{}^{\delta}$

$$h_{hk} = H_{\alpha\beta\gamma\delta}\lambda_{h|}{}^{\alpha}\lambda_{k|}{}^{\beta}\lambda_{h|}{}^{\gamma}\lambda_{k|}{}^{\delta} = r'_{hk} - r''_{hk}$$

where r_{hk}' is the Riemann curvature with respect to (2) and r_{hk}'' with respect to (1) of the orientation determined by $\lambda_{h|}$, $\lambda_{k|}$. Summing on h from 1 to m we have

$$H_k = H_{\alpha\beta\gamma\delta}a^{\alpha\gamma}\lambda_{k|}{}^{\beta}\lambda_{k|}{}^{\delta} = r_k' - r''_k \tag{9}$$

The quantity H_k depends only on the direction $\lambda_{k|}$. We will call this the *mean curvature difference* since the right side of (9) is the difference of the mean curvatures of V_m, one with respect to V_m itself and the other with respect to V_n. Summing (9) on k from 1 to m we get the well known relation [4]

$$H = R' - R''.$$

The tensor

$$H_{\beta\delta} = H_{\alpha\beta\gamma\delta}a^{\alpha\gamma} = R'_{\beta\delta} - R''_{\beta\delta} \tag{10}$$

is seen to be an invariant symmetric tensor and hence another set of principal directions are given by the solutions of

$$(H_{\beta\delta} - H_k a'_{\beta\delta})\lambda_{k|}{}^{\beta} = 0.$$

and hence we have the theorem:

The m directions through each point of V_m for which the mean curvature difference has an extreme value are mutually perpendicular.

From (10) we see that the necessary and sufficient condition that the Ricci tensor $R'_{\beta\delta}$ should coincide with the tensor $R''_{\beta\delta}$ formed with respect to V_n, is that $H_{\beta\delta}$ should vanish.

Since the invariant H contains H_k it is evident that it can be expressed as the sum of two invariants, with respect to a given direction, which in the case $m = 4$, $k = 1$ are

$$h_{12} + h_{13} + h_{14},$$
$$h_{22} + h_{24} + h_{34}.$$

These invariants depend on the direction λ_1.

4. In case $m = 3$ the tensor $R''_{\alpha\beta\gamma\delta}$ can be replaced by a

[4] Eisenhart, loc. cit., page 174. Struik, loc. cit.

tensor of the second order in the same way that $R_{\alpha\beta\gamma\delta}$ can when V_n is a space of three dimensions.[5]

$$\frac{1}{a'}R''_{2323} = b^{11}, \quad \frac{1}{a'}R''_{2313} = b^{12}, \quad \frac{1}{a'}R''_{2312} = b^{13},$$

$$\frac{1}{a'}R''_{1313} = b^{22}, \quad \frac{1}{a'}R''_{1312} = b^{23}, \quad \frac{1}{a'}R''_{1212} = b^{33}.$$

Then the principal directions of the tensor b_{ij} are the directions perpendicular to which the Riemannian curvature with respect to V_n is has an extreme value. These principal directions coincide with those determined by $R''_{\beta\delta}$. For putting

$$R''' = a'_{rs}b^{rs} = a'''_{rs}b_{rs}$$

we obtain the formula

$$R''_{\beta\delta} = R'''a'_{\beta\delta} - b_{\beta\delta}$$

and consequently the principal directions of $R''_{\beta\delta}$ and $b_{\beta\delta}$ coincide.

[5] Ricci Levi-Civita, Mathodes de calcul differentiel absolu., *Math. Ann.*, Vol. 54, page 142. Bianchi, Geometria differenziale, 2d Italian edition, Vol. I, page 353.

HERMITIAN POLYNOMIALS AND FOURIER ANALYSIS

By Norbert Wiener

It is the purpose of the present paper to bring out the relation of the development of a function in a series of Hermite orthogonal functions to its Fourier development, and to extend certain results of Hermann Weyl [1] leading to Fourier developments of fractional order.

As is well known,

$$e^{-t^2+2tz} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} t^n.$$

We wish to solve the following integral equation:

$$(1) \qquad e^{-\omega^2 t^2+2i\omega x-(x^2/2)} = \int_{-\infty}^{\infty} e^{t^2+2ity-(y^2/2)} K(x, y)dy.$$

We shall clearly have

$$\omega^n H_n(x)e^{-(x^2/2)} = \int_{-\infty}^{\infty} i^n H_n(y)e^{-(y^2/2)} K(x, y)dy,$$

so that the transformation

$$\varphi(x) = \int_{-\infty}^{\infty} K(x, y)f(y)dy$$

will replace the nth Hermite function by the same function, multiplied by the factor $(\omega/i)^n$. To solve equation (1), let us put

$$K(x, y)e^{x^2-y^2/2} = Q(x, y).$$

Then

$$\int_{-\infty}^{\infty} Q(x, y)e^{2ity}dy = e^{-t^2(\omega^2+1)+2i\omega x}.$$

[1] H. Weyl, Quantenmechanik und Gruppentheorie, *Ztsch. f. Physik*, vol. 46 (1927), pp. 1–47. Especially note pp. 22, 33.

In case $R(\omega^2 + 1) > 0$, the Fourier transformation yields us

$$Q(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-t(\omega^2+1)+2t(\omega x - ty)}dt$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-\left[t\sqrt{\omega^2+1} - \frac{\omega x - ty}{\sqrt{\omega^2+1}}\right]^2} e^{\left[\frac{\omega x - ty}{\sqrt{\omega^2+1}}\right]^2} dt$$

$$= \frac{e^{\frac{\omega^2 x^2 - 2\omega xy - y^2}{\omega^2+1}}}{\sqrt{\pi(\omega^2 + 1)}},$$

so that

$$K(x, y) = \frac{e^{(x^2+y^2)\frac{\omega^2-1}{2(\omega^2+1)} - \frac{2\omega xy}{\omega^2+1}}}{\sqrt{\pi(\omega^2 + 1)}}.$$

Let us put $\omega = e^{i\theta}$. Then

$$K(x, y) = \frac{e^{i\left[\frac{x^2+y^2}{2}\tan\theta - xy\sec\theta - \theta/2\right]}}{\sqrt{2\pi\sec\theta}}.$$

In the particular case where $\theta = 0$, this reduces to

$$K(x, y) = e^{-ixy}/\sqrt{2\pi}.$$

Let us suppose that $f(x)$ is quadratically summable over $(-\infty, \infty)$. The operator

$$\int_{-A}^{A} K(x, y)f(y)dy$$

is obtained from $f(y)$ by the following stages:

(1). We mutilate $f(y)$ into the function

$$f_A(y) = f(y), \qquad [|y| < A]$$
$$f_A(y) = 0. \qquad [|y| \geq A].$$

(2). We multiply $f_A(y)$ by $e^{i(y^2\tan\theta/2) - i(\theta/2)}$. In case θ is real, this does not affect $|f_A(y)|$.

(3). We make the Fourier transform of the function obtained from (2) as a function of z. This leaves the integral of the absolute value of the square invariant.

(4). We put $z = y\sec\theta$, and divide the whole function by $\sqrt{\sec\theta}$, to keep the integral of the square invariant.

(5). We multiply by $e^{i(x^2 \tan \theta)/2}$. If θ is real, it results from these steps that

$$\int_{-\infty}^{\infty} dx \left| \int_{-A}^{A} K(x, y) f(y) dy \right|^2 = \int_{-A}^{A} |f(y)|^2 dy$$

and that

$$\varphi(x) = \text{l. m.}_{A \to \infty} \int_{-A}^{A} K(x, y) f(y) dy$$

exists, and is such that

$$\int_{-\infty}^{\infty} |\varphi(x)|^2 dx = \int_{-\infty}^{\infty} |f(y)|^2 dy.$$

Since:

(1). The Hermite functions constitute a closed orthogonal set;

(2). The transformation

$$\varphi(x) = \text{l. m.}_{A \to \infty} \int_{-A}^{A} K(x, y) f(y) dy \tag{2}$$

is "length-preserving";

$$\omega^n H_n(x) e^{-(x^2/2)} = \int_{-\infty}^{\infty} i^n H_n(y) e^{-(y^2/2)} K(x, y) dy, \tag{3}$$

the transformation (2) simply multiplies the coefficient of $H_n(y)e^{-y^2/2}$ in the Hermite development of the quadratically summable function $f(y)$ by $(\omega/i)^n$, at least when $|\omega| = 1$. If $(\omega/i)^{2\nu} = 1$, it follows that

$$\begin{aligned} f(-x) = &\text{l. m.}_{A_1 \to \infty} \int_{-A_1}^{A_1} K(x, \xi_1) d\xi_1 \\ &\text{l. m.}_{A_2 \to \infty} \int_{-A_1}^{A_1} K(\xi_1, \xi_2) d\xi_2 \cdots \text{l. m.}_{A_n \to \infty} \int_{-A_n}^{A_n} K(\xi_{n-1}, \xi_n) f(\xi_n) d\xi_n, \end{aligned} \tag{3}$$

as the changing of the sign of the odd terms in the Hermite development of a function $f(x)$ transforms it into $f(-x)$. Where $\nu = 2$, (3) yields the ordinary Fourier repeated integral, so that it may be regarded as a generalized Fourier integral expansion. Its convergency properties—as opposed to those of convergence in the mean—seem to merit some study.

It is a matter of some interest to see how transformation (2)

acts on $iyf(y)$. We have

$$\frac{\partial K(x, y)}{\partial x} = \left[x \left(\frac{\omega^2 - 1}{\omega^2 + 1} \right) - \frac{2i\omega y}{\omega^2 + 1} \right] K(x, y).$$

Hence

$$yK(x, y) = \frac{-(\omega^2 + 1)\frac{\partial K}{\partial x} + (\omega^2 - 1)xK}{2i\omega}$$

$$= i \cos\theta \frac{\partial K}{\partial x} + x \sin\theta K.$$

It thus follows formally from (2) that

$$(4) \quad i \cos\theta \frac{d\varphi(x)}{dx} + x \sin\theta\, \varphi(x) = \underset{A\to\infty}{\text{l. m.}} \int_{-A}^{A} K(x, y) y f(y) dy.$$

In case both $f(y)$ and $yf(y)$ are quadratically summable, (4) is easily demonstrable, as we may show by actually proceeding to the limit in the mean. Thus the transformation (2) transforms the operator x into the operator $i \cos\theta(d/dx) + \sin\theta x$.

If we apply to $f(y)$ the transformation $e^{i\lambda y}$, we multiply $K(x, y)$ by $e^{i\lambda y}$. Now

$$K(x + \lambda, y) = \frac{e^{i\left[\frac{(x^2+y^2)}{2}\tan\theta + x\lambda\tan\theta + (\lambda^2/2)\tan\theta - xy\sec\theta - \lambda y\sec\theta - (\theta/2)\right]}}{\sqrt{2\pi}\sec\theta}$$

$$= K(x, y) e^{i(x\lambda\tan\theta - y\lambda\sec\theta)} e^{i(\lambda^2\tan\theta)/2}.$$

If we let $\lambda = -\mu\cos\theta$, we have

$$K(x - \mu\cos\theta, y) e^{ix\mu\sin\theta} = K(x, y) e^{iy\mu} e^{i(\mu^2\sin 2\theta)/4}.$$

Thus the transformation $e^{iy\mu}$ on $f(y)$ corresponds to the transformation which turns $\varphi(x)$ into $e^{-i[\mu x\sin\theta - (\mu^2/4)\sin 2\theta]}\varphi(x - \mu\cos\theta)$. Comparing this with our definition of the transform of $yf(y)$, we obtain

$$e^{-(\mu^2/4)i\sin 2\theta} e^{i\mu x\sin\theta} \varphi(x - \mu\cos\theta)$$

as a natural definition of

$$e^{ix\sin\theta - \cos\theta(d/dx)^2}\varphi(x).$$

This is in agreement with the definition of Weyl.

CLASSES OF FUNCTIONS, ORTHOGONAL ON AN INFINITE INTERVAL, HAVING THE POWER OF THE CONTINUUM [1]

By Philip Franklin

1. Introduction

Many of the theorems relating to the expansions of functions, defined for a finite interval, in ordinary Fourier series of sines and cosines generalize at once to other sets of orthogonal functions, complete for the interval. Such sets of orthogonal functions arise naturally as the characteristic functions of linear problems, but may be built up artificially by applying the process of orthogonalization to any complete set of functions.

The possibility of expanding functions, defined for all real numbers, in "Fourier" series of sines and cosines of all possible frequencies, and the properties of such expansions is the main theme of the theory of almost periodic functions.[2] So far as the writer knows, there are no boundary value problems which lead to alternative sets of orthogonal functions in terms of which these expansions may be made. Furthermore, the process of orthogonalization does not directly carry over, since it is a step by step process and the number of functions coming into consideration for the infinite interval is non-enumerable. Hence the question arises as to the existence of functions essentially distinct from the trigonometric functions which may be used to expand almost periodic functions.

The main object of the present note is to indicate a generalized orthogonalization process which leads to the existence of alternative basic sets. Our process requires the assumption of Zermelo's axiom.

We show that many of the relations of almost periodic functions to the sines and cosines of all frequencies remain true when these are replaced by one of our alternative basic sets.

[1] Presented to the *American Mathematical Society*, December, 1927.

[2] H. Bohr, "Zur Theorie der fastperiodischen Funktionen," I, II, III, *Acta Mathematica*, 45–47 (1924–1926).

2. Construction of the sets

We shall restrict ourselves to periodic functions, and as we shall need as periods all possible real numbers, we begin with a subdivision of the real numbers into a number of classes. With each real number a, let us associate all rational multiples of a. This leads to a non-enumerable number of classes, each containing an enumerable set of elements. Now select one number from each class, A. Then every real number is commensurable with some A, but each pair of A's has an irrational ratio. We may take $A = 1$ for the class of rational numbers.

Next consider a set of functions, complete for the unit interval $0 \leq x \leq 1$,

$$u_0, u_1, u_2, \cdots. \tag{1}$$

We take $u_0 = 1$, and define the remaining functions outside the unit interval by the equation

$$u_i(x + 1) = u_i(x), \tag{2}$$

so that they are of period 1.

Now select a particular A, and for each rational number r, define a set of functions by the equations

$$v_{Ari}(x) = u_i(x/Ar), \qquad i \geq 1. \tag{3}$$

For fixed A, the number of functions v just defined is enumerable, and we write them in order, after adding u_0 to the set, as:

$$V_0 = 1, V_1, V_2, \cdots. \tag{4}$$

These functions may be orthogonalized for the infinite interval by the usual process, provided we agree to throw out from the set any function which is linearly dependent on those which precede it. The orthogonalized functions are:

$$W_0 = 1,$$

$$W_{n+1} = \frac{V_{n+1} - \sum_{i=1}^{n} \frac{W_i}{T_n} \int_0^{T_n} W_i V_{n+1} dx}{\sqrt{\frac{1}{T_n} \int_0^{T_n} \left[V_{n+1} - \sum_{i=1}^{n} \frac{W_i}{T_n} \int_0^{T_n} W_i V_{n+1} dx \right]^2 dx}}. \tag{5}$$

The T_n appearing in this relation is the least common multiple

of the periods of the functions $V_1, V_2, \cdots V_{n+1}$. It clearly leads to the same value as the corresponding formulas with

$$\frac{1}{T_n}\int_0^{T_n} \quad \text{replaced by} \quad \lim_{T\to\infty}\frac{1}{T}\int_0^T,$$

so that the W's are normal and orthogonal for the infinite interval. T_n exists, since the periods of the V's are rationally related, and hence the W's are all periodic functions, with period a rational multiple of A.

By repetition of the procedure just outlined for all the A's, we obtain a non-enumerable set of periodic functions, such that any two coming from the same A, i.e. with commensurable periods, are orthogonal. But, we may show that two functions W_1 and W_2, each orthogonal to 1, with incommensurable periods are necessarily orthogonal. For, by a corollary of Bohr's fundamental theorem, if W_1 and W_2 are two almost periodic functions,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T W_1W_2dx = \sum_k C_1(k)C_2(k), \tag{6}$$

where the summation extends to all numbers k which are frequencies of both functions, i.e. those with non-zero coefficients $C_1(k)$, $C_2(k)$. But, in the special case at hand, since the k's are rational multiples of $2\pi/A_1$, while the k's are rational multiples of $2\pi/A_2$, the only possible common frequency is 0, and this is excluded since both functions are orthogonal to 1. Thus the right member of (6) is zero, and our contention is proved.

If we define a *basic set* of periodic functions as a non-enumerable set, normal and orthogonal for the infinite interval, and such that a sub-set of linear combinations exists which is complete for any single period, we have proved:

THEOREM I. *A process of orthogonalization exists by which a basic set may be derived from any set of continuous functions, complete for the unit interval and taking the same values at its end points.*

COROLLARY 1. *The process still leads to a basic set if a different complete set of functions be used for each period.*

For, in applying equation (3) no use was made of the fact that we used the same set u_i for each Ar. We may use distinct sets for different A's and also for different r's.

COROLLARY 2. *Every basic set, containing 1 as one of the functions, may be constructed by the process of Corollary 2.*

For, from the definition of a basic set, there are complete sets of linear combinations for every period, and if we start with these, and in (4) take the proper order, we must end up with the given set. Any function which does not at first appear in one of the complete sets, may be inserted in addition in the set corresponding to its period, which will not harm the completeness already existing.

3. *Properties of basic sets*

One of the cardinal theorems of almost-periodic functions is the possibility of approximating them by linear combinations of exponentials uniformly. In fact this is a characteristic property. That a similar property holds for any basic set, we express in

THEOREM II. *A necessary and sufficient condition for a function to be almost periodic is the existence of a sequence of linear combinations of functions $F_\lambda(x)$ of a basic set,*

$$P_n(x) = \sum_{k=0}^{N(n)} c_{nk} F_{\lambda nk}(x), \tag{7}$$

approximating it in the sense that

$$\lim_{n \to \infty} \operatorname{Max} |f(x) - P_n(x)| = 0. \tag{8}$$

As the sum of any finite number of almost periodic, or in particular, continuous periodic functions is almost periodic, properly distributed displacement numbers exist for each $P_n(x)$ for any $\epsilon/3$. On taking n so large that

$$|f(x) - P_n(x)| < \epsilon/3,$$

these displacement numbers will be displacement numbers for ϵ for $f(x)$. The uniform continuity of $f(x)$ is similarly established from that of the $P_n(x)$ so that the sufficiency of the condition is proved.

For the necessity, we recall that the theorem is known to hold when the basic set are exponentials. Let then

$$R_n(x) = \sum_{k=0}^{M(n)} g_{nk} e^{-\Omega_{nk} x}, \tag{9}$$

where

$$|f(x) - R_n(x)| < \epsilon/2. \tag{10}$$

From the completeness of the $F_\lambda(x)$, the kth exponential of (9) may be approximated uniformly by a linear combination of F's to within $\epsilon/2|g_{nk}|M$, and the sum of these terms multiplied by g_{nk} will give a $P_n(x)$ uniformly approximating $f(x)$ to within ϵ.

We write

$$M\{f(x)\} = \lim_{T\to\infty} \int_0^T f(x)dx \tag{11}$$

to indicate mean values, and define coefficients with respect to any set of basic functions by

$$A_\lambda = M\{f(x)F_\lambda(x)\}. \tag{12}$$

With these abbreviations, we may express the analogue of the fundamental theorem as

THEOREM III. *There are only an enumerable number of functions of any basic set for which a given almost periodic function has non-vanishing coefficients. For these coefficients we have the completeness relation*

$$M\{f(x)^2\} = \sum_{k=1}^{\infty} A_{\lambda k}^2. \tag{13}$$

We prove this by considering the mean square error of an arbitrary linear combination. From the fact that our basic functions are normal and orthogonal, we have:

$$M\left\{\left|f(x) - \sum_{k=1}^{N} c_k F_{\lambda_k}(x)\right|^2\right\} = M\{f^2\} - 2\sum_{k=1}^{N} c_k A_{\lambda_k} + \sum_{k=1}^{N} c_k^2 \tag{14}$$

$$= M\{f^2\} - \sum_{k=1}^{N} A_{\lambda_k}^2 + \sum_{k=1}^{N} (A_{\lambda_k} - c_k)^2.$$

This shows that for given N, the mean square error is least when the c_k are equal to the A_{λ_k}, the N largest being chosen. Furthermore, from the positive nature of the squares, and means of squares,

$$\sum_{k=1}^{N} A_{\lambda_k}^2 \leq M\{f^2\} \tag{15}$$

so that the series on the right of (13) converges absolutely. Also

$$M\{f^2\} - \sum_{k=1}^{N} A_{\lambda_k}^2 \leq M\left\{\left|f - \sum_{k=1}^{N} c_k F_{\lambda_k}\right|^2\right\}. \tag{16}$$

By the approximation theorem, c_k exist for which the right member of (16) is arbitrarily small, and hence the left member can be made small, and in the limit, $N = \infty$, (15) reduces to the equality (13).

Practically all the corollaries of the fundamental theorem, e.g., the unique determination of an almost periodic function by its coefficients, hold for basic sets. The proofs are identical with those in the original theory.

4. Discontinuous basic functions. Generalized almost periodic functions

If we allow our original functions to be piecewise continuous, the process of section 2 may still be carried through to yield a normal and orthogonal set. The only place where we have used the continuity of our functions is in deriving (6) as a property of almost periodic functions. This equation is proved by applying the fundamental theorem to $(W_1 + W_2)$, W_1 and W_2, so that it holds for any type of functions closed with respect to addition, for which the fundamental theorem holds. In particular, the pseudo-periodic functions [1] are such a class, and they include periodic functions with simple discontinuities.

For these discontinuous basic sets, Theorem I, with its corollaries, the part of Theorem II which deals with the necessity of the condition, and Theorem III may be proved as before.

The approximation theorem for almost periodic functions in terms of exponentials, of which Theorem II above is the analogue for basic functions, has analogues for the various types of generalized almost periodic functions which have been defined.[2] Most of these theorems are true for any basic sets with the appropriate type of continuity; and may be proved by using the possibility of approximating the exponentials by the basic functions, and vica versa, in the sense used for the particular type of generalized function under discussion.

[1] Cf. N. Wiener, *Math Zeitschr.*, 24 (1925), pp. 575–616; or W. Stepanoff, *Math. Annalen*, 95 (1926), pp. 472–498, for a treatment of these functions.

[2] Cf. the author's "Approximation theorems for generalized almost periodic functions," *Math. Zeitschr.*, 29 (1928), pp. 70–86.

A NEW METHOD IN THE THEORY OF QUANTICS. SECOND PAPER

By Frank L. Hitchcock

Introduction. In Art. 1–6 a concept of invariants is developed which is more general than that of the classical theory in the sense that the quantities concerned satisfy some of the ordinary laws of algebra but need not (although they may) satisfy the rest. The relation of the method to the usual symbolic method is established. The discussion is limited to rational homogeneous invariants with cogredient indices under an arbitrary linear transformation. In the remainder of the paper a method is given for expressing any such invariant as a sum of determinants.

1. **Résumé of the logical foundations of the method.**

In part I (this Journal, vol. IV, p. 238, July 1925) were assumed n linearly independent symbols $\mathbf{e}_1, \cdots, \mathbf{e}_n$ called **basis vectors** in terms of which any other vector $\mathbf{a}$ is linearly expressible,

$$\mathbf{a}_{nn} = a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n. \tag{1}$$

The **components** $a_1, \cdots, a_n$ are commutative and associative in addition and multiplication, their negatives exist, i.e. $a_i - a_i = 0$, and multiplication is distributive with respect to addition; their reciprocals are not necessary. No arithmetical nor geometrical properties are (in general) attributed either to components or to vectors. Integers enter the scheme by abbreviating $a + a = 2a$, etc., but as coefficients merely.

Addition of vectors, when it occurs, is associative and commutative. Multiplication of components with vectors, and of vectors with each other, is distributive with respect to addition. Multiplication of vectors with each other is associative but (in general) not commutative.

Comparing with the postulates of algebra as stated for example by E. V. Huntington in "The Fundamental Postulates of Alge-

bra" (Monographs on Modern Mathematics, p. 187) it will be noted that none of postulates 12–27 is assumed.

It is very easy to see that the present system resembles that laid down by H. Weyl in "*Raum, Zeit, Materie*" (5 Aufl. p. 15) where, however, components are the reals of ordinary algebra, and vectors are interpreted in affine space. The use of non-commutative symbols to express invariants goes back, in embryo at least, to Hamilton [1] and Cayley.[2] The latter was the first to employ ordered determinants and determinants of higher class.[3] The intimate connection between the Aronhold symbolic method and the Gibbs polyad was first utilized by H. B. Phillips.[4] Much use of this connection has been made by J. A. Schouten and D. J. Struik.[5] The expression of the **space complement** as an ordered determinant is due to A. Naess.[6] The importance of the algebra underlying geometrical systems is now universally recognized (e.g. the algebra of tensors). E. Study [7] has even said "Mag man die Vektoranalysis um gewisser Anwendungen willen oder als selbstständige Disziplin betreiben, immer handelt es sich dabei zunächst um Gegenstände der Algebra, und zwar um invarianten gewisser Gruppen linearen Transformationen,"—whereas the founders of vector algebra sought a calculus of the space elements themselves. A. Whitehead remarks [8] "In some connexions it would be better to abandon the conception of originals studied by the aid of substitutive schemes, and to conceive of two sets of interrelated

[1] Hamilton, *Elements*, Art. 352.

[2] Cayley, *Collected Papers*, No. 20, Vol. 1, p. 123; Phil. Mag., Vol. XXVI (1845), p. 141.

[3] Cayley, *Collected Papers*, No. 12, Vol. 1, p. 63; Trans. Cam. Phil. Soc., Vol. VIII, pp. 1–16.

[4] H. B. Phillips, *Some Invariants and Covariants of Ternary Collineations*, Am. Jour. Math., **36** (1914), p. 311.

[5] J. A. Schouten and D. J. Struik, *Einführung in die Neueren Methoden der Differentialgeometrie.*

D. J. Struik, *Grundzüge der mehrdimensionalen Differentialgeometrie in Direkter Darstellung.*

[6] A. Naess, *On a special polyadic of order n–p which can be derived from any p independent vectors in an n-dimensional space and which can be regarded as a generalisation of the vector product.* Videnskapsselkapets Skrifter I Mat.-naturv. Klasse, 1922, No. 13.

[7] E. Study, *Einleitung in die Theorie der Invariantern linearen Transformationen auf Grund der Vektorrechnung*, S. 3.

[8] A. Whitehead, *Universal Algebra*, p. 12.

things studied together, each scheme exemplifying the operation of the same general laws."

The most general element employed in the present system is the **polyadic**: a product of p vectors is a p-ad (**polyad**) and a sum of p-ads is a p-adic. Thus any polyadic $\mathbf{A}$ can be written

$$\mathbf{A} = \sum_{j=1}^{h} \mathbf{a}_{1j}\mathbf{a}_{2j} \cdots \mathbf{a}_{pj} \tag{2}$$

where

$$\mathbf{a}_{kj} = a_{kj\alpha}\mathbf{e}_\alpha \quad \text{(summed on } \alpha = 1, 2, \cdots, n\text{)}. \tag{3}$$

By expressing all vector factors as in (3) and collecting terms the polyadic may be written

$$\mathbf{A} = A_{\alpha_1 \cdots \alpha_p}\mathbf{e}_{\alpha_1} \cdots \mathbf{e}_{\alpha_p} \quad \text{(summed on all indices)}. \tag{4}$$

Thus to every polyadic and a set of basis vectors there corresponds a matrix of n^p components given by the important relation

$$A_{i_1 \cdots i_p} = \sum_{j=1}^{h} a_{1ji_1} \cdots a_{pji_p}. \tag{5}$$

Two polyadics (or two vectors) are equal when and only when corresponding components are equal each to each. A polyadic or a vector vanishes only when all components vanish. Hence the product of two polyadics is zero only when one of the polyadics vanishes in all components.

If the basis vectors undergo linear transformation $\mathbf{T}$ the components of a vector $\mathbf{e}_\alpha a_\alpha$ will be transformed by

$$a_i' = a_\alpha T_{\alpha i}. \tag{6}$$

The transform of $\mathbf{a}$ will be denoted by $\mathbf{a}'$. When all vector factors on the right of (2) are transformed the new polyadic will be denoted by $\mathbf{A}'$.

A polynomial in any set of components is defined to be a sum of terms wherein each term is a product of these components with, it may be, a coefficient not belonging to the set. Only homogeneous polynomials will be considered.

Two polynomials can be equal (for the present) only when they are homogeneous of the same degree in the same set of components and have corresponding coefficients equal each to each.

A polynomial is zero only when all coefficients are zero. Components do not in general have *values* in an arithmetic sense; they merely satisfy the laws above laid down.

If, however, in the polyadic (4), the basis vectors of the p respective indices be replaced by p distinct sets of variables

$$x_{i_1}^{(1)}, \cdots, x_{i_p}^{(p)}$$

i.e. variables in the sense of ordinary algebra, we obtain a polynomial linear in each set, a p-linear n-ary **form**

$$A_{\alpha_1 \cdots \alpha_p} x_{\alpha_1}^{(1)} \cdots x_{\alpha_p}^{(p)}. \tag{7}$$

Other forms are obtainable by causing some of the p sets of variables to be no longer distinct.

The following definition of an invariant is in form the same as the classical one (for cogredient variables), differing from it by implying the above more general concept of components or coefficients and by including invariants of polyadics, matrices, and forms. Let F be a polynomial homogeneous in the components of a set of polyadics **A**, **B**, $\cdots$, **G**. For brevity we may write

$$F = F(A_{i_1 \cdots i_p}, \cdots, G_{k_1 \cdots k_t}) \tag{8}$$

where $p, \cdots, t$ are the classes of the respective matrices and where $A_{i_1 \cdots i_p}$, etc., denote not individual components but entire sets; or still more briefly

$$F = F(\mathbf{A}, \cdots, \mathbf{G}). \tag{8a}$$

Let F' be the same polynomial in the transformed components, i.e.

$$F' = F(\mathbf{A}', \cdots, \mathbf{G}'). \tag{9}$$

If it be true that

$$F' = \Delta^k F \tag{10}$$

where Δ is the determinant of the n^2 quantities $T_{\alpha i}$, the polynomial F is an invariant of the given set of polyadics and of the related matrices and forms.

2. Slugs.

Consider now the product of k determinants made up in the following manner. Let a polyad be the product of kn^p vector

factors. Divide these factors into k groups of n vectors each in any manner, with a definite order among the vectors of each group, which need not be the same as the original order in the polyad. Form the determinant of the n^2 components of the vectors in each group, the order of rows being the same as the order of vectors in the group. Multiply together the k determinants thus formed.

These products, the stuff that invariants are made of, are basic to the argument. They are of **weight** k. Any such product will be called a **slug** of the polyad.

If the k determinant factors of a slug are written out and multiplied together the resulting polynomial, linear in the components of each of the kn vector factors of the polyad, will contain $(n!)^k$ terms, no two of which are in general like terms. (Even if two vector factors be equal they are distinguished by their positions in the polyad.)

Definition I. That term of a slug which is the product of the elements of the main diagonals of the determinant factors will be called the **leading term** of the slug.

That is, if the vector factors of the polyad, after grouping, be temporarily denoted by

$$\mathbf{a}_1\mathbf{a}_2 \cdots \mathbf{a}_n,\ \mathbf{b}_1\mathbf{b}_2 \cdots \mathbf{b}_n,\ \cdots,\ \mathbf{g}_1\mathbf{g}_2 \cdots \mathbf{g}_n \qquad (11)$$

then the leading term of the slug

$$(\mathbf{a}_1\mathbf{a}_2 \cdots \mathbf{a}_n)(\mathbf{b}_1\mathbf{b}_2 \cdots \mathbf{b}_n) \cdots (\mathbf{g}_1\mathbf{g}_2 \cdots \mathbf{g}_n) \qquad (12)$$

is

$$a_{11}a_{12} \cdots a_{nn}b_{11}b_{22} \cdots b_{nn} \cdots g_{11}g_{22} \cdots g_{nn}. \qquad (13)$$

With this notation any other term of the same slug may, aside from sign, be obtained from the leading term by permuting the values of the second subscript, i.e. performing a substitution on these digits. The substitution is *restricted* in the following manner: it may affect a single group, such as $a_{11}a_{22} \cdots a_{nn}$, alone, or may be compounded of substitutions affecting several groups, but it must not shift digits from group to group. Each term has of course a proper sign.

A slug is evidently a special case of a polynomial linear in the

components of the given polyad. Such a polynomial is in general the sum of n^{kn} terms of the form

$$a_{ii_1}a_{si_2}\cdots a_{pi_p} \tag{14}$$

where $p = kn$. Each term may carry a coefficient not a function of the polyad.

Definition II. Any term of the form (14) in which precisely k of the indices $i_1 \cdots i_p$ have the value *1*, and precisely k have the value *2*, and so on, each value from *1* to n occurring precisely k times, will be called **equipollent.**

It is evident by (13) that all terms of a slug are equipollent. Conversely any equipollent term with coefficient $+1$ is the leading term of a certain number of slugs. This number is the number of distinct slugs obtainable by grouping the kn vectors of (11), *without altering subscripts*, so that in each group the subscript runs in order from *1* to n, namely $(k!)^{n-1}$ ways. The fact may be stated as

Theorem I. *Any equipollent term with coefficient* $+1$ *is the leading term of* $(k!)^{n-1}$ *slugs.*

3. **Reversion coefficients.**

The sum of the $(k!)^{n-1}$ slugs having a common leading term will be called the *reversion polynomial* of that term, or simply its **reversion.** It will be shown later that if the reversion polynomials of all the terms of a given slug be added (each with the sign carried by its own term in the slug) we recover the original slug multiplied by a function of k and n called the **reversion function.**

The coefficient of any term of a reversion polynomial will be called the **reversion coefficient** of that term. (This of course assumes that in forming a reversion like terms have been collected.) A reversion coefficient is thus by definition an integer (positive, negative, or zero) relating a component of the original polyad, taken as leading term, to a second component. (The reversion coefficient of the leading term with respect to itself is obviously $(k!)^{n-1}$.) It is in fact the algebraic sum of the number of ways by which it is possible to pass from one component to

the other using the groupings and substitutions mentioned in Art. 2. It may be easily shown that the relation is a reciprocal one: if C_1 and C_2 are two components of the polyad the reversion coefficient of C_2 with respect to C_1 is the same as that of C_1 with respect to C_2; in other words the two components yield each a reversion polynomial, such that the coefficient of C_2 in the first is the same as the coefficient of C_1 in the second.

I shall now illustrate slugs, reversions, reversion coefficients, and the reversion function in the simple case $k = n = 2$. Let the given polyad be **abcd**. There are six distinct slugs, three being negatives of the other three. $\left(\text{In general } \frac{2(kn)!}{k(k!)(n!)} \text{ distinct slugs.}\right)$ There are $\frac{(kn)!}{(k!)^n}$ equipollent components or six in this case,

$$a_1b_1c_2d_2 \quad a_1b_2c_2d_1 \quad a_2b_1c_1d_2 \quad a_2b_2c_1d_1 \quad a_1b_2c_1d_2 \quad a_2b_1c_2d_1$$

The component $a_1b_1c_2d_2$ is leading term of the two slugs **(ac)(bd)** and **(ad)(bc)** that is of

$$(a_1c_2 - a_2c_1)(b_1d_2 - b_2d_1) = a_1b_1c_2d_2 - a_1b_2c_2d_1 - a_2b_1c_1d_2 + a_2b_2c_1d_1$$

and of

$$(a_1d_2 - a_2d_1)(b_1c_2 - b_2c_1) = a_1b_1c_2d_2 - a_1b_2c_1d_2 - a_2b_1c_2d_1 + a_2b_2c_1d_1.$$

The reversion of $a_1b_1c_2d_2$ is the sum of these two slugs. The reversions of the six equipollent components may be represented by the matrix of the reversion coefficients,

	$a_1b_1c_2d_2$	$a_1b_2c_2d_1$	$a_2b_1c_1d_2$	$a_2b_2c_1d_1$	$a_1b_2c_1d_2$	$a_2b_1c_2d_1$
$a_1b_1c_2d_2$	+ 2	− 1	− 1	+ 2	− 1	− 1
$a_1b_2c_2d_1$	− 1	+ 2	+ 2	− 1	− 1	− 1
$a_2b_1c_1d_2$	− 1	+ 2	+ 2	− 1	− 1	− 1
$a_2b_2c_1d_1$	+ 2	− 1	− 1	+ 2	− 1	− 1
$a_1b_2c_1d_2$	− 1	− 1	− 1	− 1	+ 2	+ 2
$a_2b_1c_2d_1$	− 1	− 1	− 1	− 1	+ 2	+ 2

symmetrical by the reciprocal relation above pointed out. The first four rows are the reversions of the terms of **(ab)(cd)**; changing the signs of the second and third rows and omitting

the fifth and sixths rows and then adding we recover six times this slug. Thus the reversion function has the value 6 for this case. The six reversions are equal in pairs. (In general equal in groups of $n!$.) It can be shown that any reversion polynomial can be written as a determinant.

4 The transform of an equipollent component.

Any term of the form (14), transformed by (6), becomes

$$a_{\beta\alpha_1}T_{\alpha_1 i_1}\cdots a_{\beta\alpha_p}T_{\alpha_p i_p}, \tag{15}$$

a polynomial of n^p terms, all unlike. Each term of this transform is a component of the original polyad multiplied by p elements of the transformation $T_{\alpha i}$. I shall suppose the untransformed term (14) to be equipollent. Some of the terms of the transform will also be equipollent, and among them will be groups of terms containing the *same* p elements $T_{\alpha i}$.

For example let $k = n = 2$ with original polyad **abcd**. Choose as untransformed term (14) the component $a_1b_1c_2d_2$ (alphabetical order replacing the first index for brevity). Of the 16 terms of the transform six are equipollent and may be arranged as

$$a_1b_1c_2d_2T_{11}^2T_{22}^2 + (a_1b_2c_1d_2 + a_1b_2c_2d_1 + a_2b_1c_1d_2 + a_2b_1c_2d_1)T_{11}T_{12}T_{21}T_{22} + a_2b_2c_1d_1T_{12}^2T_{21}^2. \tag{16}$$

That is the four components in parentheses have for coefficient the same four elements of the transformation.

Definition III. Given a polyad of kn vector factors and a chosen equipollent component, a group of equipollent components which, in the transform, are multiplied by the same kn elements of the transformation will be said to constitute a **species.**

Theorem II. *Any coefficient of a species is, aside from numerical factor, a term in the expansion of* Δ^k.

Proof. The values of the *second* index in the kn elements of the transformation are the same as those which occur in the untransformed component since $a_i' = a_\alpha T_{\alpha i}$, etc.; hence each value occurs k times, the original term being equipollent. The values of the *first* index also occur k times each for they are the same as those of the terms of the species. But the terms of Δ^k consist

of the totality of all products of kn elements $T_{\alpha i}$ equipollent on both indices. The theorem therefore holds.

It is important to examine in precisely what manner the components of a species differ from one another. As just pointed out there are, in the coefficient of the species, k of the $T_{\alpha i}$ whose second indices have the value unity because they have operated to transform k letters originally carrying the value unity. In passing from one component of the species to another the occurrence of these k letters is altered only in the manner in which the *first* indices of the same k $T_{\alpha i}$ are distributed among them, and all possible distributions occur: such a distribution is necessary and sufficient, so far as these k letters are concerned, that the components belong to the species. A similar remark applies to each of the n sets of k letters each. In other words *a component* C_1 *can always be carried into any other* C_2 *of the same species by a substitution which interchanges only letters whose indices were originally equal in numerical value.* Conversely any such substitution applied to a component produces a component of the same species (perhaps the same component).

For example the four components in parentheses in (16) pass into one another by interchanging a with b, or c with d, or both.

Theorem III. *All components of a species have the same reversion coefficient with respect to the untransformed term.*

Proof. Let the restricted substitutions of Art. 2 by which a term of a slug is produced from its leading term be denoted by R. Call the untransformed component C_0 and take it as leading term. Let the substitutions above mentioned which interchange only letters having, in the leading term, indices of the same value, be denoted by S. As defined in Art. 2, a component occurs in the reversion polynomial by reason of a substitution S followed by a substitution R. The various occurrences of C_1 may be expressed by equations, $C_1 = C_0S_1R_1$, $C_1 = C_0S_2R_2$, etc. Let S_{12} denote a substitution which carries C_1 into C_2 as above shown. Then $C_2 = C_0S_1R_1S_{12}$, $C_2 = C_0S_2R_2S_{12}$, etc. But since the R's and the S's operate on different symbols,—the R's on the indices, the S's on the letters, or, if we prefer the notation (14), on the first subscript,—any R is commutative with any S.

Thus $C_2 = C_0S_1S_{12}R_1$, $C_2 = C_0S_2S_{12}R_2$, etc. Now S_1S_{12}, S_2S_{12}, etc. are themselves substitutions S. We accordingly have the set of substitutions $S_1S_{12}R_1$, $S_2S_{12}R_2$, etc. which produce C_2 corresponding in order to the set S_1R_1, S_2R_2, etc. which produce C_1. Furthermore, since any two of the second set are distinct by definition, the corresponding pair of the first set will be distinct; for corresponding substitutions differ only in the factor S_{12} which possesses a reciprocal.

Also, corresponding substitutions have the same sign; for the S's carry no sign and the R is common.

By similar reasoning, to each SR producing C_2 may be made to correspond, by means of S_{21}, one and only one SR producing C_1; and it will have the same sign.

Hence the correspondence is 1 : 1; and the algebraic sum of the number of positive and negative substitutions is the same for both components. But these two sums are $k!$ times the respective reversion coefficients; for the S are $(k!)^n$ in number and the sum of all the C_0SR is $k!$ times the reversion polynomial. (We may suppose, for example, that the reversion is conducted by omitting all S which derange the letters bearing in (11) the subscript unity; then each of the $k!$ derangements of these letters merely reproduces the reversion.) Hence the reversion coefficients of C_1 and C_2 are equal as was to be proved.

The common reversion coefficient of the terms of a species may be called the **reversion coefficient of the species.**

Corollary I. *If, in the transform of a chosen equipollent term, each species be multiplied by its reversion coefficient and the results added, the sum is the reversion polynomial of the chosen term.*

If two or more chosen equipollent terms are transformed, those species, in each case, which multiply the same combination of the T_{ai} may be called **like species.**

Corollary II. *Like species, each with respect to its own leading term, have the same reversion coefficient.*

Corollary III. *If an invariant F is linear in the components of a polyad, it can be written as a linear function of its slugs.*

For the equation $F' = \Delta^k F$ is equivalent, by comparing like combinations of the T_{ai}, to a system of equations: on the left of

each is a linear function of like species, on the right F multiplied by an integer (+, −, or zero). Multiplying by the reversion coefficients and adding, the left side of the sum is a linear function of reversions, i.e. of sums of slugs; the right is ξF where ξ is an integer, the **reversion function.** Hence the corollary. (It can be shown that ξ is equal to the sum of the squares of all the reversion coefficients divided by $(k!)^{n-1}$.)

Corollary IV. *If the reversions of the terms of a slug be added, each with the sign of its own term, the sum is ξ times the original slug.*

For a slug is evidently an invariant, linear in the components of the polyad, to which the process of Corollary III may be applied.

Corollary V. *An invariant linear in the components of a polyad contains only equipollent terms.*

Corollary VI. *Only polyads with kn vector factors possess linear invariants.*

5. **Linear invariants.**

The abbreviated term **linear invariant** will be used to denote an invariant linear in the components of a given polyad or polyadic. The following lemmas scarcely require formal proof.

Lemma I. A linear invariant of a polyad is an invariant of its vector factors; and is linear in the components of each of them.

Lemma II. Conversely an invariant linear in the components of each of a set of vectors is a linear invariant of any polyad of which these vectors are the factors.

Lemma III. An invariant of a polyad $\mathbf{a}_1 \cdots \mathbf{a}_p$ is an invariant of the p linear forms $a_{1\alpha_1}x_{\alpha_1}{}^{(1)}, \cdots, a_{p\alpha_p}x_{\alpha_p}{}^{(p)}$ and conversely.

A linear invariant of a polyadic is evidently a sum of linear invariants of polyad terms, and the functional symbol F is distributive over the polyads; that is, if a polyadic $\mathbf{P}$ be developed

$$\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2 + \cdots + \mathbf{P}_h \tag{17}$$

where the terms $\mathbf{P}_1, \cdots, \mathbf{P}_h$ are polyads, a linear invariant $F(\mathbf{P})$ may be developed

$$F(\mathbf{P}) = F(\mathbf{P}_1) + F(\mathbf{P}_2) + \cdots + F(\mathbf{P}_h). \tag{18}$$

I shall now show that any invariant whatever, in the sense defined in Art. 1, is a linear invariant of a properly chosen polyadic. Let the invariant be of degrees $\alpha, \beta, \cdots, \eta$ in the components of $\mathbf{A}, \mathbf{B}, \cdots, \mathbf{G}$ respectively. Let $\mathbf{A}^\alpha$ denote the product of α polyadics all equal to $\mathbf{A}$ and similarly for all cases.

Definition IV. A polyadic $\mathbf{P}$ defined by

$$\mathbf{P} = \mathbf{A}^\alpha \mathbf{B}^\beta \cdots \mathbf{G}^\eta \tag{18}$$

will be called a **source** of the given invariant.

Theorem IV. *An invariant is a linear invariant of its source.*

Proof. By hypothesis the invariant is a sum of terms each of which is of degrees $\alpha, \cdots, \eta$ in the components of $\mathbf{A}, \cdots, \mathbf{G}$. But the components of a source are precisely the totality of products of this form. Hence the invariant is a linear function of the components of the source. Moreover an invariant of $\mathbf{A}, \cdots, \mathbf{G}$ is an invariant of $\mathbf{P}$ for, by definition, the polyadics $\mathbf{A}, \cdots, \mathbf{G}$ and $\mathbf{P}$ are transformed by transforming their vector factors, i.e.

$$\mathbf{P}' = \mathbf{A}'^\alpha \cdots \mathbf{G}'^\eta \tag{19}$$

hence the components of $\mathbf{P}'$ are the same functions of the components of $\mathbf{A}', \cdots, \mathbf{G}'$ that those of $\mathbf{P}$ were of $\mathbf{A}, \cdots, \mathbf{G}$. Therefore the transformed invariant F' is the same function of $\mathbf{P}'$ that F was of $\mathbf{P}$. The theorem is therefore proved.

Corollary. *Any invariant is a linear function of slugs.*

For by (18) the invariant may be expressed as a sum of linear invariants of polyads, and by Corollary III to Theorem III, as a linear function of the slugs of these polyads.

6. **The symbolic method.**

The relationship between the present method and the usual symbolic method may now be briefly stated. The transition is made by virtue of the following assumptions.

I. A polyadic may be represented by, and is symbolically equal to, any one of its polyad terms. Thus (2) is equivalent to the symbolic equations

$$\mathbf{A} = \mathbf{a}_{11}\mathbf{a}_{21} \cdots \mathbf{a}_{p1} = \mathbf{a}_{12}\mathbf{a}_{22} \cdots \mathbf{a}_{p2} = \cdots = \mathbf{a}_{1h}\mathbf{a}_{2h} \cdots \mathbf{a}_{ph} \tag{20}$$

which by taking components are equivalent to

$$\mathbf{A}_{i_1 \cdots i_p} = \mathbf{a}_{1 i_1} \mathbf{a}_{2 i_2} \cdots \mathbf{a}_{p i_p} = \cdots = \mathbf{a}_{1 h_1} \mathbf{a}_{2 h_2} \cdots \mathbf{a}_{p h_p}. \quad (21)$$

If we pass from polyadics to forms by (7) the symbolic equations (20) become the usual representation of a form by products of linear forms wherein products of several different sets of p linear forms are taken as equivalent.

II. A polyadic $\mathbf{A}^\alpha$ is symbolically equal to the product of α polyads all different. Thus in the symbolic sense

$$\mathbf{A}^\alpha = \mathbf{a}_{11} \mathbf{a}_{21} \cdots \mathbf{a}_{p1} \mathbf{a}_{12} \mathbf{a}_{22} \cdots \mathbf{a}_{p2} \cdots \mathbf{a}_{1\alpha} \mathbf{a}_{2\alpha} \cdots \mathbf{a}_{p\alpha} \quad (22)$$

and similarly in all cases.

It follows that a source **P** as defined by (18) is symbolically equal to a single polyad which is the product of several polyads similar to (22), one for each of the given polyadics **A**, ···, **G**. Such a polyad I shall call a **symbolic source.**

Assuming each of the polyadics **A**, ···, **G** written as the sum of not less than α, ···, η polyads, respectively, a symbolic source is in form the same as one of the actual terms in the development of the actual source.

III. A slug of a symbolic source is symbolically equal to a particular invariant of **A**, ···, **G**.

From a symbolic slug we may pass to the actual invariant denoted by it in two ways,—

A. Sum the slug over the terms of the actual source. That is, let the source be developed as in (17), denote the slug by F, and sum as in (18). The particular invariant denoted by the slug is thus obtained as a sum of actual slugs.

B. Multiply out the determinants whose product constitutes the slug, and substitute actual for symbolic components by (22). This is the method of the classical theory.

To show that the classical theory is logically contained in the present theory it is now merely necessary to prove

Theorem V. *The processes* (A) *and* (B) *are equivalent.*

Proof. Let the classes of the matrices of **A**, ···, **G** (i.e. the number of vector factors in each polyad) be respectively p_a, ···, p_g. A slug is the product of k determinants, so that multiplying out yields $(n!)^k$ terms. Consider one of these terms: it

contains αp_a components of the vector factors of **A**, namely α sets of p_a components each; in each set the second subscript has the same value for all components (notation of (21)), while the first subscript runs from *1* to p_a; the third subscript or index has, for each component, a particular value, as $i_1, i_2, \cdots, i_p$. In method (B) such a set is at once replaced by $A_{i_1 \cdots i_p}$.

On the other hand method (A) implies the expansion of $\mathbf{A}^\alpha$, that is a set of α summations of the form (2). There is thus exactly one summation (on the second subscript) for every set of p_a components above mentioned, and by (5) this summation gives $A_{i_1 \cdots i_p}$. Similarly the components of **B**, $\cdots$, **G** obtained by the two methods are the same, hence they are equivalent.

The corollary to Theorem IV may be stated symbolically as *Any invariant may be denoted symbolically by a linear function of symbolic slugs.* This is the **first fundamental theorem** for cogredient variables.

Remarks on the above argument. The basic proposition is theorem III. The third corollary, by virtue of Lemma III, is equivalent to the theorem, basic in the classical theory, that *an invariant of linear forms is an aggregate of determinants.*

Of other arguments to the same end, using a series development,[9] seminvariants,[10] or the Cayley operator Ω,[11] the latter is closest to that employed in the present paper: it is known that Ω^k applied to (14) yields an aggregate of determinants; I have found no investigation of the character of this aggregate.

If however we note that the operation of taking a derivative is equivalent to a sum of substitutions,[12] we see that applying Ω^k to (14) is equivalent to performing on (11) the totality of the substitutions S and adding the corresponding slugs. The resulting aggregate of determinants is therefore $k!$ times the reversion of the untransformed term. To say it another way, a first application of Ω yields k^n terms each of which contains one of the determinants whose main diagonals make up the leading term; continuing the process to the end, each resulting slug has this leading term and occurs $k!$ times.

[9] Clebsch, *Th. d. Bin. Formen*, Art. 9.
[10] H. Andoyer, *Leçons sur la theorie des formes*, Chap. III.
[11] Grace and Young, *Alg. of Inv.*, Chap. II.
[12] Turnbull, *Theory of Determinants, Matrices, and Invariants*, p. 119.

The reversion function ξ may now be easily evaluated. Let

$$\Omega^k F' = cF \tag{23}$$

where it is known [13] that

$$c = n! \frac{(n+1)!}{1!} \cdots \frac{(n+k-1)!}{(k-1)!}. \tag{24}$$

If F is a slug we therefore have by theorem III, corollary IV, $k!\xi F = cF$, hence

$$k!\xi = c. \tag{25}$$

More generally, any concomitant, no matter how formulated, is actually a sum of permutations. Since the same statement may be made of determinants of general class whose elements are also determinants, it is to be expected that some, if not all, concomitants can be expressed in determinant form,—as was pointed out long ago by Von Escherich [14] for certain cases. (A second paper contemplated by him was never published.) For some investigations a symbolic expression, or a summation, suffices, but not for the study of particular invariants: even in comparatively simple cases methods (A) and (B) above, by which an invariant can, in principle, be explicitly developed, are practically unworkable.

In part I it was shown that the process of **folding** [15] always yields a concomitant and also affords the means of expressing this concomitant as a determinant; while at the same time it presents us with a symbolic form of the concomitant which, by the reasoning of Art. 6 (applicable with slight change to concomitants in general) is equivalent to the ordinary symbolic form. In the next article I shall consider one phase of the converse problem: *given a symbolic expression for an invariant, to interpret it as a folding operation and hence to develop the invariant by aid of determinants.*

[13] Weitzenbock, *Invariantentheorie*, S. 16.

[14] Von Escherich, Denks. d. K. Ak. d. W. Wien., **43**, II (1882), pp. 1–12. *Die Det. hoheren Ranges und ihre Verwendung zur Bildung von Invarianten.*

[15] The folding process is a development of a method due to Lepine Hall Rice, *Determinantal Expression of Multiple Cross and Dot Products of Polyadics in three Dimensions*, this Journal, Vol. IV, No. 2 (April 1925), p. 130.

7. **Determinantal formulation of invariants.**

I shall show that any invariant given by a product of symbolic determinants (symbolic slug) can be written as a sum of determinants. It will first be necessary to define the folding process in a manner slightly more general than already done (*A Theory of Ordered Determinants*, etc., this Journal, Vol. IV, No. 4, July 1925, pp. 205–237, to be referred to as OD).

Let there be q polyadics ($q \leqq n$) $\mathbf{A}, \cdots, \mathbf{G}$ and their corresponding matrices $A_{i_1 \cdots i_a}, \cdots, G_{m_1 \cdots m_g}$, where $a, \cdots, g$ are the classes of these matrices. Take the product of the q polyadics, and let it be **multiple** on the first r indices of every polyadic but **direct** on the remaining indices (OD p. 212). Replace each group of q first factors, q second factors, etc. up to each group of q r^{th} factors by its space complement, i.e. fold the q polyadics on their first r factors. The remaining factors remain unaltered in every polyadic. The result is independent in *fact* of the particular manner in which the polyadics were written as sums of polyads, i.e. it is **distributively invariant** (Part I, p. 240). But in *form* the result may be written as a determinant by supposing the polyadics developed in the following manner: expand each of the first r factors of every polyadic in terms of basis vectors and collect terms; the first polyadic $\mathbf{A}$ then takes the form

$$\mathbf{A} = \mathbf{e}_{\alpha_1}\mathbf{e}_{\alpha_2} \cdots \mathbf{e}_{\alpha_r}\mathbf{M}_{\alpha_1\alpha_2\cdots\alpha_r}, \tag{26}$$

the $\mathbf{M}$'s being $(a-r)$-adics; replace the last $a-r$ indices by a single multipartite index running from *1* to n^{a-r}, i.e. treat the $\mathbf{M}$'s as vectors in space of n^{a-r} dimensions. If we set

$$\mathbf{E}_u = \mathbf{e}_{i_{r+1}}\mathbf{e}_{i_{r+2}} \cdots \mathbf{e}_{i_a} \tag{27}$$

$\mathbf{A}$ may be written

$$\mathbf{A} = \mathbf{e}_{\alpha_1}\mathbf{e}_{\alpha_2} \cdots \mathbf{e}_{\alpha_r}\mathbf{E}_u A_{\alpha_1\alpha_2\cdots\alpha_r u} \tag{28}$$

where u is the (a–r)-partite index, summed from *1* to n^{a-r}. Expanding the other polyadics $\mathbf{B}, \cdots, \mathbf{G}$ similarly, with multipartite indices $v, \cdots, z$, the result of folding is a determinant of $r+1$ directions constructed as follows:

1°. The first $n-q$ layers are alike and have r-ad elements, precisely as in OD theorem VII.

2°. The remaining q layers consist, layer by layer, of the respective sets of multipartite vectors $\mathbf{E}_u\mathrm{A}_{\alpha_1\ldots\alpha_r u}$, etc., arising from expansions like (28).

3°. In developing the determinant the order in the first $n-q$ layers is multiple, i.e. all first factors precede all second factors, etc. Otherwise the order of all factors is that of the original polyadics.

4°. The determinant is of full sign. If $r+1$ is odd the first direction, i.e. that of the layers as above defined, is non-signant.

The proof is identical with that given in this Journal, Vol. IV, No. 3, p. 179, by Mr. L. H. Rice and the author, scalar components of that proof being replaced by multipartite vectors; for the proof is independent of the commutativity of the scalar components.

If the integers a, $\cdots$, g are equal we recover the case treated in OD, except for direct development now taken on indices after the r^{th} of every polyadic.

To take first an example involving only two-way determinants, let $n=3$, $q=2$, $r=1$, $a=3$, $b=2$, i.e. fold a triadic into a dyadic in three-space on the first factor only. Write

$$\mathbf{A} = \mathbf{e}_{\alpha_1}\mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}\mathrm{A}_{\alpha_1\alpha_2\alpha_3} = \mathbf{e}_\alpha\mathbf{M}_\alpha \tag{29}$$

where the $\mathbf{M}$'s are dyadics, and

$$\mathbf{B} = \mathbf{e}_{\beta_1}\mathbf{e}_{\beta_2}\mathrm{B}_{\beta_1\beta_2} = \mathbf{e}_\beta\mathbf{b}_\beta \tag{30}$$

where the $\mathbf{b}$'s are vectors. The folding, denoted by $[\mathbf{A}, \mathbf{B}]^1$, gives the square determinant

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \mathbf{M}_1 & \mathbf{M}_2 & \mathbf{M}_3 \\ \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 \end{vmatrix} \tag{31}$$

in which the development is to be direct, with rows ordinal. By putting for the $\mathbf{M}$'s and the $\mathbf{b}$'s their values we have

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}\mathrm{A}_{1\alpha_2\alpha_3} & \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}\mathrm{A}_{2\alpha_2\alpha_3} & \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}\mathrm{A}_{3\alpha_2\alpha_3} \\ \mathbf{e}_{\beta_2}\mathrm{B}_{1\beta_2} & \mathbf{e}_{\beta_2}\mathrm{B}_{2\beta_2} & \mathbf{e}_{\beta_2}\mathrm{B}_{3\beta_2} \end{vmatrix} \tag{32}$$

which develops into the tetradic

$$\mathbf{e}_1\mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}\mathbf{e}_{\beta_2}(A_{1\alpha_2\alpha_3}B_{3\beta_2} - A_{3\alpha_2\alpha_3}B_{1\beta_2}) + \cdots + \cdots. \quad (33)$$

If we prefer matrices to polyadics we may indicate (33) by letting ϵ_{rst} be a cubic matrix whose elements are $+1$ or -1 when formed from ϵ_{123} by positive or negative permutation respectively, and otherwise zero; the components of (33) are then

$$\epsilon_{r\alpha_1\beta_1}A_{\alpha_1 i_2 i_3}B_{\beta_1 j_2}. \quad (33A)$$

If, on the other hand, we are dealing with algebraic forms, we may suppose the basis vectors $\mathbf{e}_{\alpha_2}$, $\mathbf{e}_{\alpha_3}$, and $\mathbf{e}_{\beta_2}$ to denote three sets of cogredient variables, while the basis vectors of the first row in (32) denote contragredient variables: then (33) is a concomitant of a trilinear form and a bilinear form. If the three sets of cogredient variables are the same we have a concomitant of a ternary cubic and a quadratic.

To take an example leading to cubic determinants, let the above triadic and dyadic be folded on their first two factors. The result, denoted by $[\mathbf{A}, \mathbf{B}]^2$, may be written

$$\begin{vmatrix} \mathbf{e}_1\mathbf{e}_1, & \mathbf{e}_1\mathbf{e}_2, & \mathbf{e}_1\mathbf{e}_3 \\ \mathbf{e}_2\mathbf{e}_1, & \mathbf{e}_2\mathbf{e}_2, & \mathbf{e}_2\mathbf{e}_3 \\ \mathbf{e}_3\mathbf{e}_1, & \mathbf{e}_3\mathbf{e}_2, & \mathbf{e}_3\mathbf{e}_3 \end{vmatrix} \begin{vmatrix} \mathbf{e}_{\alpha_3}A_{11\alpha_3}, & \mathbf{e}_{\alpha_3}A_{12\alpha_3}, & \mathbf{e}_{\alpha_3}A_{13\alpha_3} \\ \mathbf{e}_{\alpha_3}A_{21\alpha_3}, & \mathbf{e}_{\alpha_3}A_{22\alpha_3}, & \mathbf{e}_{\alpha_3}A_{23\alpha_3} \\ \mathbf{e}_{\alpha_3}A_{31\alpha_3}, & \mathbf{e}_{\alpha_3}A_{32\alpha_3}, & \mathbf{e}_{\alpha_3}A_{33\alpha_3} \end{vmatrix} \begin{vmatrix} B_{11}, & B_{12}, & B_{13} \\ B_{21}, & B_{22}, & B_{23} \\ B_{31}, & B_{32}, & B_{33} \end{vmatrix} \quad (34)$$

and develops into a triadic which equals

$$\mathbf{e}_1\mathbf{e}_1(\mathbf{e}_{\alpha_3}\mathbf{A}_{22\alpha_3}B_{33} - \mathbf{e}_{\alpha_3}A_{23\alpha_3}B_{32} - \mathbf{e}_{\alpha_3}A_{32\alpha_3}B_{23} + \mathbf{e}_{\alpha_3}A_{33\alpha_3}B_{22}) \quad (35)$$

plus eight similar expressions. In matrix language (34) is

$$\epsilon_{r\alpha_1\beta_1}\epsilon_{s\alpha_2\beta_2}A_{\alpha_1\alpha_2 i_3}B_{\beta_1\beta_2}. \quad (36)$$

The first two factors of (35), corresponding to r and s of (36), have contravariant character.

The case $n = 3$, $q = 2$, $r = 2$, $a = 3$, $b = 4$, i.e. folding a triadic and a tetradic in 3-space on two factors, gives a cubic determinant in form like OD (28) but the β's are dyadics.

The case $n = 3$, $q = 3$, $r = 2$, $a = 3$, $b = 4$, $c = 4$, i.e. folding a triadic and two tetradics in 3-space on two factors gives a cubic determinant in form like OD (27) but the β's and γ's are dyadics.

The case $n = 4$, $q = 2$, $r = 2$, $a = 3$, $b = 4$ gives a cubic determinant like OD (31) but the φ's are vectors instead of dyadics.

It is interesting to notice that the "general term" of the development, although more general than (38) and (39) of OD, can be more simply written, by virtue of the direct development on the indices after the r^{th} of every polyadic, and the consequent possibility of employing multipartite indices. With the notation of OD we replace the indices $r_1 \cdots s_1$ by u; $r_2 \cdots s_2$ by v; up to $r_q \cdots s_q$ by z, where $u, \cdots, z$ are multipartite as above described. Everything that concerns the first r indices of any polyadic is precisely as in OD. Thus any non-vanishing component of $[\mathbf{A}, \cdots, \mathbf{G}]^r$, that is of the polyadic obtained by folding any q polyadics on their first r indices or vector factors, will multiply a polyadic

$$\mathbf{E}^{(1)}_{g_1 \cdots g_{n-q}} \cdots \mathbf{E}^{(r)}_{g_1 \cdots g_{n-q}} \mathbf{E}^{(1)}_{u} \cdots \mathbf{E}^{(q)}_{z} \tag{37}$$

and the component will itself be a determinant of $r + 1$ directions and of order q whose layers of the first direction have the respective elements

$$\mathrm{A}_{\mathrm{A} \cdots \mathrm{Gu}}, \ \mathrm{B}_{\mathrm{A} \cdots \mathrm{Gv}}, \ \cdots, \ \mathrm{G}_{\mathrm{A} \cdots \mathrm{Gz}} \tag{38}$$

wherein the indices A, $\cdots$, G (r in number) are respectively complementary to the a's, $\cdots$, g's of (37), i.e. run through the values not possessed by the latter,—according to a Laplace expansion. The values of the multipartite indices are *fixed* throughout any one component. Each component carries also a sign, determined precisely as in OD, since it concerns only the folded indices; the sign is merely a matter of Laplace development. Finally, we may express $[\mathbf{A}, \cdots, \mathbf{G}]^r$ in matrix language as

$$\epsilon^{(1)}_{r_1 \cdots r_{n-q} \alpha_1 \cdots \eta_1} \cdots \epsilon^{(r)}_{t_1 \cdots t_{n-q} \alpha_r \cdots \eta_r} \mathrm{A}_{\alpha_1 \cdots \alpha_r i_{r+1} \cdots i_a} \cdots \mathrm{G}_{\eta_1 \cdots \eta_r m_{r+1} \cdots m_g}. \tag{39}$$

The expansion (28) is equivalent to arranging the components of $\mathbf{A}$ in a rectangular $(r + 1)$-way block instead of in a "cubical" a-way block. There will be n elements in each file of the first r directions but n^{a-r} elements in a file of the multipartite direction u. When $\mathbf{A}$ is folded with $q - 1$ other polyadics any

component of the resulting new polyadic $[\mathbf{A}, \cdots, \mathbf{G}]^r$, that is any of the expressions (39), may be constructed by first picking out a layer of the multipartite direction, i.e. giving a fixed value to u in $\mathbf{M}_{i_1 \ldots i_r u}$; we then pick from this layer (which is an r-way "cube") a particular q-rowed block as given by the Laplace development, and have as result the first layer of the required component: the other $q - 1$ layers are picked in like manner from $\mathbf{B}, \cdots, \mathbf{G}$. Since each value of u can occur with each value of $v, \cdots, z$ we shall have in general $n^{(a+b+\cdots+g-q^2)}$ distinct non-vanishing components for *each* term in the Laplace development. For example the expression which is explicitly written in (33) is the first of three terms in developing (32), and contains $3^{(3+2-2\times1)} = 27$ components, as is evident since it involves summation on three indices, α_2, α_3, and β_2. Again, (35) involves only $3^{(3+2-2\times2)} = 3$ components, i.e. summation on α_3.

It may be remarked that there is no logical reason for confining folding to the *first* r indices of every polyadic. But there is a gain in simplicity of statement. If folding is on *any* r indices we may usually suppose an isomer or conjugate of the polyadic to be taken such that this isomer is folded as described above.

I now assume an invariant to be indicated by a product of symbolic determinants. Let the invariant be of degree α in the components of $\mathbf{A}$, and denote $\mathbf{A}$ symbolically by

$$\mathbf{A} = \mathfrak{a}_1\mathfrak{a}_2 \cdots \mathfrak{a}_{p_a} = \mathfrak{a}_1'\mathfrak{a}_2' \cdots \mathfrak{a}_{p_a}' = \cdots = \mathfrak{a}_1^{(\alpha)}\mathfrak{a}_2^{(\alpha)} \cdots \mathfrak{a}_{p_a}^{(\alpha)} \quad (39)$$

and similarly for $\mathbf{B}, \cdots, \mathbf{G}$, the other "ground-polyadics". Out of the k symbolic determinants whose product indicates the invariant let that determinant be selected which contains symbols of the greatest number of the polyadics $\mathbf{A}, \cdots, \mathbf{G}$; and call this number q. (If $q < n$ it is obvious that symbols of at least one polyadic occur twice or more in the determinant). We then fold these q polyadics on a set of q indices or vector factors (one for each polyadic) whose symbols occur in the selected determinant; if, as above remarked, these are not the respective first factors of the polyadics we can bring them into first place in every case by isomerization. If it happens that symbols of the *same* q polyadics occur also in another symbolic determinant, so much the better: we fold on *two* sets of q indices each (selected

ad libitum if two or more symbols of the same polyadic occur in one determinant; only one symbol being taken for each polyadic from any one determinant). Similarly if symbols of all q polyadics occur in a third determinant we fold on three sets of q factors each, and so on: If r is the total number of determinants in which all q of these polyadics occur we isomerize so that the folding is on the first r indices or vector factors as described above. Call the result $\mathbf{L}_1$.

The remaining polyadics are treated in the same way, but with reference only to the symbolic determinants not involved in the first operation. Call the result of folding the largest possible number of these remaining polyadics on the largest possible number of indices $\mathbf{L}_2$. Proceeding in this way we shall either exhaust the ground polyadics or there will remain a single polyadic which may be called $\mathbf{L}$. We thus have a set of new polyadics $\mathbf{L}_1, \mathbf{L}_2, \cdots, \mathbf{L}$ fewer in number than $\mathbf{A}, \cdots, \mathbf{G}$, the components of which (excepting perhaps $\mathbf{L}$) are determinants as proved above, since they are formed by folding.

We now treat $\mathbf{L}_1, \cdots, \mathbf{L}$ as $\mathbf{A}, \cdots, \mathbf{G}$ were treated. If there is more than one polyadic in the series $\mathbf{L}_1, \cdots, \mathbf{L}$, and unless the given invariant is merely a product of two invariants of lower degree,—a case here to be regarded as trivial,—there will certainly remain one or more symbolic determinants not so far involved in the selection and containing factors occurring in two or more of the polyadics $\mathbf{L}_1, \cdots, \mathbf{L}$. Since these symbols have not yet been employed it is obvious they are not yet folded. We accordingly fold $\mathbf{L}_1, \cdots, \mathbf{L}$ by sets containing as many polyadics as possible, by the same rule used to form $\mathbf{L}_1, \cdots, \mathbf{L}$ from $\mathbf{A}, \cdots, \mathbf{G}$. Call the new set $\mathbf{M}_1, \cdots, \mathbf{M}$, where, it may be, $\mathbf{M}$ is one of the original polyadics $\mathbf{A}, \cdots, \mathbf{G}$ unchanged or one of the series $\mathbf{L}_1, \cdots, \mathbf{L}$ unchanged. Continuing in this manner we shall finally arrive at a result in which all the given polyadics have been included in one folding operation.

This result is the same as if we had begun with a symbolic source (Art. 5) and had introduced various space complements in place of groups of factors, but by the steps employed we have the result in form of a determinant which may be called a **folded source** and denoted by $\mathbf{D}$. It may be that $\mathbf{D}$ is precisely the

required invariant, namely when it agrees in symbolic form with the given symbolic invariant. But in general some of the folding operations by which **D** was built up will involve fewer than n polyadics, i.e. we shall have space complements of fewer than n vectors. It follows that **D** will in general be a polyadic, a concomitant of the given polyadics.

The required invariant is finally to be obtained from **D** by replacing any set of n vector factors, some of which may already have been replaced by their space complements, by the space complement, i.e. determinant, of all of them, as indicated by the given symbolic invariant. This is equivalent to forming an algebraic sum of properly selected components of **D**. Since these components are determinants we have obtained the invariant as an algebraic sum of determinants.

Take, for an extremely simple example, a vector $\mathbf{a}$ and a dyadic $\mathbf{B}$ in 3-space. Let $\mathbf{B} = \mathbf{b}_1\mathbf{b}_2$ symbolically. Required the invariant $(\mathbf{ab}_1\mathbf{b}_2)$ as a sum of determinants. First construct $(\mathbf{ab}_1)\mathbf{b}_2$, i.e. fold the vector with the first factor of the dyadic; yielding the determinant

$$\mathbf{D} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ \mathbf{e}_\beta B_{1\beta} & \mathbf{e}_\beta B_{2\beta} & \mathbf{e}_\beta B_{3\beta} \end{vmatrix}, \tag{40}$$

a dyadic, a concomitant of $\mathbf{a}$ and $\mathbf{B}$, contravariant on the prefactors and covariant on the postfactors; the nine components are the elements of the matrix

$$\begin{matrix} a_2B_{31} - a_3B_{21} & a_2B_{32} - a_3B_{22} & a_2B_{33} - a_3B_{23} \\ a_3B_{11} - a_1B_{31} & a_3B_{12} - a_1B_{32} & a_3B_{13} - a_1B_{33}. \\ a_1B_{21} - a_2B_{11} & a_1B_{22} - a_2B_{12} & a_1B_{23} - a_2B_{13} \end{matrix} \tag{41}$$

We form $(\mathbf{ab}_1\mathbf{b}_2)$ from $(\mathbf{ab}_1)\mathbf{b}_2$ by retaining in (40) only the values of β equal to the subscript of the contravariant prefactors by which $\mathbf{e}_\beta$ is multiplied,—as is seen directly from the definition of a space complement. The required invariant is therefore the sum of the elements of the main diagonal of (41).

We can, if we prefer, write the steps in matrix language. The symbolic determinant $(\mathbf{ab}_1\mathbf{b}_2)$ is replaced by

$$\epsilon_{rst}a_rB_{st}; \tag{42}$$

the matrix (41) is the same as

$$\epsilon_{rst}a_sB_{tu} \tag{43}$$

and the passage from $(\mathbf{ab}_1)\mathbf{b}_2$ to $(\mathbf{ab}_1\mathbf{b}_2)$ appears as the identity

$$\epsilon_{1st}a_sB_{t1} + \epsilon_{2st}a_sB_{t2} + \epsilon_{3st}a_sB_{t3} = \epsilon_{rst}a_sB_{tr} = \epsilon_{rst}a_rB_{st}. \tag{44}$$

As an example in which a folded source **D** (or better D) is identical with the required invariant, take three binary dyadics $\mathbf{A} = a_1a_2$, $\mathbf{B} = \mathbf{b}_1\mathbf{b}_2$, and $\mathbf{C} = \mathbf{c}_1\mathbf{c}_2$, symbolically. Required to develop the invariant $(\mathbf{a}_1\mathbf{b}_1)(\mathbf{a}_2\mathbf{c}_1)(\mathbf{b}_2\mathbf{c}_2)$. Form first $(\mathbf{a}_1\mathbf{b}_1)\mathbf{a}_2\mathbf{b}_2$, i.e. fold **A** into **B** on the first factor only; giving the determinant

$$\begin{vmatrix} \mathbf{e}_\alpha A_{1\alpha} & \mathbf{e}_\alpha A_{2\alpha} \\ \mathbf{e}_\beta B_{1\beta} & \mathbf{e}_\beta B_{2\beta} \end{vmatrix} \tag{45}$$

a dyadic with components $A_{1\alpha}B_{2\beta} - A_{2\alpha}B_{1\beta}$ $(\alpha, \beta = 1, 2)$. We now form the required invariant at one step by folding this dyadic into **C** on both factors since evidently $(\mathbf{a}_1\mathbf{b}_1)\mathbf{a}_2\mathbf{b}_2$ thus folded with $\mathbf{c}_1\mathbf{c}_2$ gives $(\mathbf{a}_1\mathbf{b}_1)(\mathbf{a}_2\mathbf{c}_1)(\mathbf{b}_2\mathbf{c}_2)$. By the rules given above it is the cubic determinant

$$\begin{vmatrix} A_{11}B_{21} - A_{21}B_{11} & A_{11}B_{22} - A_{21}B_{12} \\ A_{12}B_{21} - A_{22}B_{11} & A_{12}B_{22} - A_{22}B_{22} \end{vmatrix} \begin{vmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{vmatrix}. \tag{46}$$

This is of course an invariant of three bilinear forms corresponding to the three given binary dyadics. If the dyadics or bilinear forms are all symmetrical it is well known in the classical theory [16] that $(\mathbf{ab})(\mathbf{ac})(\mathbf{bc})$ is the determinant

$$\begin{vmatrix} A_{11} & A_{12} & A_{22} \\ B_{11} & B_{12} & B_{22} \\ C_{11} & C_{12} & C_{22} \end{vmatrix}. \tag{47}$$

It follows that (46) and (47) are equal in case $A_{12} = A_{21}$, $B_{12} = B_{21}$, and $C_{12} = C_{21}$, as may be verified by expansion.

When $n = 2$ a folded source can always be found identical with an invariant given by a single product of symbolic determinants, i.e. such an invariant can always be expressed as a single determinant; because when $n = 2$ the above method calls for no folding with fewer than n symbols.

To take a less simple example, let four ternary triadics be $\mathbf{A} = \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$, $\mathbf{B} = \mathbf{b}_1\mathbf{b}_2\mathbf{b}_3$, $\mathbf{C} = \mathbf{c}_1\mathbf{c}_2\mathbf{c}_3$, and $\mathbf{G} = \mathbf{g}_1\mathbf{g}_2\mathbf{g}_3$, symbolically. Required the invariant $(\mathbf{a}_1\mathbf{b}_1\mathbf{c}_1)(\mathbf{a}_2\mathbf{b}_2\mathbf{g}_1)(\mathbf{a}_3\mathbf{c}_2\mathbf{g}_2)(\mathbf{b}_3\mathbf{c}_3\mathbf{g}_3)$ as a sum of determinants. No two symbolic determinants contain symbols of the same three triadics. Perform first the folding indicated by $(\mathbf{a}_1\mathbf{b}_1\mathbf{c}_1)$, i.e. fold **ABC** on the first factor only yielding symbolically

$$(\mathbf{a}_1\mathbf{b}_1\mathbf{c}_1)\mathbf{a}_2\mathbf{a}_3\mathbf{b}_2\mathbf{b}_3\mathbf{c}_2\mathbf{c}_3 \tag{48}$$

keeping for the moment direct order on the factors not folded. This gives a covariant of **A**, **B**, **C**, namely the determinant

$$\begin{vmatrix} \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}A_{1\alpha_2\alpha_3} & \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}A_{2\alpha_2\alpha_3} & \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}A_{3\alpha_2\alpha_3} \\ \mathbf{e}_{\beta_2}\mathbf{e}_{\beta_3}B_{1\beta_2\beta_3} & \mathbf{e}_{\beta_2}\mathbf{e}_{\beta_3}B_{2\beta_2\beta_3} & \mathbf{e}_{\beta_2}\mathbf{e}_{\beta_3}B_{3\beta_2\beta_3} \\ \mathbf{e}_{\gamma_2}\mathbf{e}_{\gamma_3}C_{1\gamma_2\gamma_3} & \mathbf{e}_{\gamma_2}\mathbf{e}_{\gamma_3}C_{2\gamma_2\gamma_3} & \mathbf{e}_{\gamma_2}\mathbf{e}_{\gamma_3}C_{3\gamma_2\gamma_2} \end{vmatrix} \tag{49}$$

a hexadic whose 729 components are the square determinants of the third order

$$\begin{vmatrix} A_{1\alpha_2\alpha_3} & A_{2\alpha_2\alpha_3} & A_{3\alpha_2\alpha_3} \\ B_{1\beta_2\beta_3} & B_{2\beta_2\beta_3} & B_{3\beta_2\beta_3} \\ C_{1\gamma_2\gamma_3} & C_{2\gamma_2\gamma_3} & C_{3\gamma_2\gamma_3} \end{vmatrix} \tag{50}$$

where the six Greek indices have the values 1, 2, 3 each with each. In matrix notation (48) is

$$\epsilon_{r3l}A_{ri_2i_3}B_{3j_2j_3}C_{lk_2k_3}. \tag{51}$$

It will be convenient temporarily to denote the determinants (50) by $L_{\alpha_2\alpha_3\beta_2\beta_3\gamma_2\gamma_3}$ so that the expansion of (49) is

$$\mathbf{L} = \mathbf{e}_{\alpha_2}\mathbf{e}_{\alpha_3}\mathbf{e}_{\beta_2}\mathbf{e}_{\beta_3}\mathbf{e}_{\gamma_2}\mathbf{e}_{\gamma_3}L_{\alpha_2\alpha_3\beta_2\beta_3\gamma_2\gamma_3}. \tag{52}$$

We now form the symbolic isomer of (48)

$$(\mathbf{a}_1\mathbf{b}_1\mathbf{c}_1)\mathbf{a}_2\mathbf{c}_2\mathbf{b}_3\mathbf{b}_2\mathbf{a}_3\mathbf{c}_3 \tag{53}$$

guided by the fact that the first three factors $\mathbf{a}_2\mathbf{c}_2\mathbf{b}_3$ are to be folded with $\mathbf{g}_1\mathbf{g}_2\mathbf{g}_3$ in order and similarly for $\mathbf{b}_2\mathbf{a}_3\mathbf{c}_3$. The actual isomer may be called $\bar{\mathbf{L}}$ and is

$$\bar{\mathbf{L}} = \mathbf{e}_{\alpha_2}\mathbf{e}_{\gamma_2}\mathbf{e}_{\beta_3}\mathbf{e}_{\beta_2}\mathbf{e}_{\alpha_3}\mathbf{e}_{\gamma_3}L_{\alpha_2\alpha_3\beta_2\beta_3\gamma_2\gamma_3}. \tag{54}$$

We next fold $\bar{\mathbf{L}}$ into **G**, the first three vector factors of $\bar{\mathbf{L}}$ with the three vector factors of **G** respectively, yielding symbolically

$$\mathbf{D} = (\mathbf{a}_1\mathbf{b}_1\mathbf{c}_1)(\mathbf{a}_2\mathbf{g}_1)(\mathbf{c}_2\mathbf{g}_2)(\mathbf{b}_2\mathbf{g}_3)\mathbf{b}_2\mathbf{a}_3\mathbf{c}_3 \qquad (55)$$

which, since all four given triadics are included, is what was above called a folded source. It may, as shown, be written in actual form as a 4-way determinant of the third order: the elements of the first layer are contravariant triads $\mathbf{e}_a\mathbf{e}_b\mathbf{e}_c$; of the second layer triadics $\mathbf{e}_{\beta_2}\mathbf{e}_{\alpha_3}\mathbf{e}_{\gamma_3}\mathrm{L}_{A\alpha_3\beta_2GB\gamma_3}$; and of the third layer $\mathrm{G}_{A'B'C'}$. **D** is a concomitant of **A**, **B**, **C**, **G**, a hexadic whose components are four-way determinants of the second order. The easiest way now to form the required invariant is to conceive $\mathbf{e}_a\mathbf{e}_b\mathbf{e}_c$ as a tripartite vector $\mathbf{E}^u$ in 27-space and $\mathbf{e}_{\beta_2}\mathbf{e}_{\alpha_3}\mathbf{e}_{\gamma_3}$ similarly as $\mathbf{E}_v$ giving **D** as a dyadic (double triadic, i.e. prefactors and postfactors triadics) in 27-space. Replacing symbolically $(\mathbf{a}_2\mathbf{g}_1)\mathbf{b}_2$ by $(\mathbf{a}_2\mathbf{g}_1\mathbf{b}_2)$ is equivalent to retaining only those values of the index β_2 which are equal to the index of $\mathbf{e}_a$ in the same term of **D**, and similarly for the other two remaining Greek indices. This is the same as saying that we add the components of **D** in which $u = v$, i.e. the elements of the main diagonal of the matrix of **D** when written as a dyadic in 27 space. We have thus the required invariant as the sum of 27 4-way determinants of the second order

$$\epsilon_{aAA'}\epsilon_{bBB'}\epsilon_{cCC'}\mathrm{L}_{Ab\alpha CB\gamma}\mathrm{G}_{A'B'C'} = \epsilon_{1AA'}\epsilon_{1BB'}\epsilon_{1CC'}\mathrm{L}_{A11CB1}\mathrm{G}_{A'B'C'}$$

plus 26 similar expressions, $(a, b, c = 1, 2, 3)$.

THE FOURIER TRANSFORM THEOREM

By Andrew C. Berry

Introduction

The Plancherel theory of the Fourier transform is one of the cornerstones of modern harmonic analysis, in that not only its theorems, but also the methods used in establishing them, find frequent application in the problems of harmonic analysis. The present paper was written to furnish the author an introduction to the vectorial methods of n-dimensional harmonic analysis. The paper is based on the n-dimensional form of the Riesz-Fischer theory of mean convergence, the fundamental theorem of which is stated without proof.

1. **Preliminary Calculations.** Let P denote, indiscriminately, either an arbitrary point of a real n-dimensional space, or the vector extending from a fixed origin to this point. Let $f(P)$ be any integrable,[1] complex function of P. By

$$\int_V f(P)dV_P$$

will be meant the n-dimensional volume integral of $f(P)$ extended over the region[2] V. In particular,

$$\int_{|P-P_0|\leq R} f(P)dV_P$$

will be the volume integral of $f(P)$ over an n-dimensional "sphere" whose radius is R and whose center is at $P = P_0$. If we denote by $(r, \phi_1, \phi_2, \cdots, \phi_{n-2}, \theta)$ a system of n-dimensional spherical

[1] Throughout this paper "integrable" will mean "integrable in the sense of Lebesgue"; all integrals that appear will be Lebesgue integrals. This implies, among other things, that the functions in question are absolutely integrable. Cf. H. Lebesgue: "Leçons sur l'intégration et la recherche des fonctions primitives," (Paris, Gauthier-Villars, 1904).

[2] Throughout this paper "region" will be used as if synonymous with "measurable n-dimensional point set."

coördinates with pole at the origin, then

$$\int_{|P|\leq R} f(P)dV_P = \int_0^R r^{n-1}dr \int_0^\pi \sin^{n-2}\phi_1 d\phi_1 \int_0^\pi \sin^{n-3}\phi_2 d\phi_2 \cdots \int_0^\pi \sin\phi_{n-2}d\phi_{n-2}\int_0^{2\pi} f(r, \phi_1, \phi_2, \cdots, \phi_{n-2}, \theta)d\theta.$$

A relation useful in the evaluation of such an iterated integral is

$$(1)\qquad \prod_{k=1}^{\nu}\left\{\int_0^\pi \sin^k\phi d\phi\right\} = \prod_{k=1}^{\nu}\left\{\sqrt{\pi}\,\frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}+1\right)}\right\} = \frac{\pi^{\nu/2}}{\Gamma\left(\frac{\nu}{2}+1\right)}.$$

The following relations concerning Bessel functions will also be useful:[3]

$$(2)\qquad J_m(x) = \frac{x^m}{2^m\sqrt{\pi}\,\Gamma\left(m+\frac{1}{2}\right)}\int_0^\pi \sin^{2m}\phi e^{ix\cos\phi}d\phi.$$

$$(3)\qquad \frac{d}{dx}[x^m J_m(x)] = x^m J_{m-1}(x).$$

$$(4)\qquad \lim_{s\to\infty}\frac{a^m}{b^{m-1}}\int_0^s J_m(ax)J_{m-1}(bx)dx = \begin{cases}1, \text{ if } b<a;\\ 0, \text{ if } b>a;\end{cases}\quad \begin{pmatrix}a>0,\\ b>0,\\ m>0\end{pmatrix}.$$

$$(5)\qquad \frac{A^m}{B^m}\int_0^\infty \frac{J_m(Ax)J_m(Bx)}{x}dx = \frac{1}{2m} \text{ if } 0<B\leqq A;\ m>0.$$

Denoting the scalar product of two real n-dimensional vectors P and Q by $(P\cdot Q)$ we shall verify the three following facts:

$$(\mathrm{I})\qquad \frac{1}{(2\pi)^{n/2}}\int_{|P|\leq R} e^{i(P\cdot Q)}dV_P = \frac{R^{n/2}J_{n/2}(R|Q|)}{|Q|^{n/2}}.$$

$$(\mathrm{II})\qquad \lim_{s\to\infty}\left(\frac{R}{2\pi}\right)^{n/2}\int_{|Q|\leq S}\frac{J_{n/2}(R|Q|)}{|Q|^{n/2}}e^{i(P\cdot Q)}dV_Q = \begin{cases}1, \text{ if } |P|<R;\\ 0, \text{ if } |P|>R;\end{cases}\quad \begin{pmatrix}|P|>0,\\ R>0\end{pmatrix}.$$

[3] Cf. Watson: "The theory of Bessel functions," pp. 18, 25, 405, 406.

$$\text{(III)}\quad \frac{\Gamma\left(\frac{n}{2}+1\right)A^{n/2}}{\pi^{n/2}B^{n/2}}\int_{\omega}\frac{J_{n/2}(A|Q+K|)J_{n/2}(B|Q+K|)}{|Q|^{n}}\,dV_Q$$
$$= 1, \text{ for any fixed } K, \text{ if } 0 < B \leqq A.$$

The integral in (I), being extended over a sphere, depends only on the magnitude and not on the direction of the vector Q. Hence, in evaluating this integral we may assume Q to be directed along the ray $\phi_1 = 0$. Thus, using (1), (2) and (3), we find:

$$\frac{1}{(2\pi)^{n/2}}\int_{|P|\leq R} e^{i(P\cdot Q)}dV_P = \frac{1}{(2\pi)^{n/2}}\int_{|P|\leq R} e^{i|Q|r\cos\phi_1}dV_P$$
$$= \int_0^R r^{n-1}\frac{J_{(n/2)-1}(|Q|r)}{(|Q|r)^{(n/2)-1}}\,dr = \frac{R^{n/2}J_{n/2}(R|Q|)}{|Q|^{n/2}}.$$

Next, since

$$\left(\frac{R}{2\pi}\right)^{n/2}\int_{|Q|\leq S}\frac{J_{n/2}(R|Q|)}{|Q|^{n/2}}e^{i(P\cdot Q)}dV_Q$$
$$= \frac{R^{n/2}}{|P|^{(n/2)-1}}\int_0^S J_{n/2}(Rq)J_{(n/2)-1}(|P|q)dq,$$

we see that (II) is a direct consequence of (4). Finally, since [4]

$$|J_m(x)| < \frac{C}{\sqrt{x}},\ x \text{ sufficiently large, } m \geqq 0,$$

it follows that the infinite integral (III) is absolutely convergent, and hence that:

$$\frac{\Gamma\left(\frac{n}{2}+1\right)A^{n/2}}{\pi^{n/2}B^{n/2}}\int_{\omega}\frac{J_{n/2}(A|Q+L|)J_{n/2}(B|Q+L|)}{|Q+L|^{n}}\,dV_Q$$
$$= \frac{\Gamma\left(\frac{n}{2}+1\right)A^{n/2}}{\pi^{n/2}B^{n/2}}\int_{\omega}\frac{J_{n/2}(A|Q|)J_{n/2}(B|Q|)}{|Q|^{n}}\,dV_Q$$
$$= \frac{\Gamma\left(\frac{n}{2}+1\right)A^{n/2}}{\pi^{n/2}B^{n/2}}\cdot\frac{2(\pi)^{n/2}}{\Gamma\left(\frac{n}{2}\right)}\cdot\int_0^{\infty}\frac{J_{n/2}(Aq)J_{n/2}(Bq)}{q}\,dq.$$

[4] Watson, loc. cit., p. 199.

Since $\Gamma\left(\frac{n}{2}+1\right)=\frac{n}{2}\Gamma\left(\frac{n}{2}\right)$ we see that (III) follows from (5).

One further fact that will be useful later is that the volume of a sphere of radius R is

$$\int_{|P|\leq R} dV_P = \frac{\pi^{n/2}R^n}{\Gamma\left(\frac{n}{2}+1\right)}.$$

2. **The Fourier Transform Theorem.** In studying the behavior of an n-dimensional volume integral as the region of integration is allowed to expand indefinitely we shall require that this expansion take place in accordance with the two following restrictions:

(i) Corresponding to any stage of expansion there must exist a sphere of finite radius, center at the origin, which will contain in its interior all points of the given region at the given stage of expansion.

(ii) Corresponding to almost any point P (i.e. any point not on a certain set of zero n-dimensional measure) there must exist a stage of expansion such that for all later stages the region will contain P in its interior.

We shall write that a function is q.i. over V ("quadratically integrable" over V), [or q.i. over ∞], if the function is integrable over any sub-region of V, [or over any finite region], and is also of integrable (absolute) square over V, [or over the entire n-dimensional space].

We state, without proof, the n-dimensional form of the RIESZ-FISCHER THEOREM.[5]

If $f_\lambda(P)$ is q.i. over ∞ for all values of the parameter λ, and if

$$\lim_{\substack{\lambda\to\infty\\ \mu\to\infty}} \int_\infty |f_\lambda(P)-f_\mu(P)|^2 dV_P = 0$$

where $\lambda, \mu \to \infty$ in any fashion, then there exists a function $f(P)$,

[5] The theory of mean convergence is discussed by Plancherel and by Titchmarsh in the papers referred to later. Other references are: E. Fischer, *Comptes Rendus*, 148 (1907), pp. 1022 and 1148. F. Riesz, *Comptes Rendus*, 148 (1907), p. 615; *Math. Annalen*, 69 (1910), p. 449. H. Weyl, *Math. Annalen*, 68 (1909), p. 225. W. H. and G. C. Young, *Quarterly Journal*, 44 (1913), p. 49.

defined uniquely almost everywhere, q.i. over ∞, and such that

$$\lim_{\lambda \to \infty} \int_{\infty} |f(P) - f_\lambda(P)|^2 dV_P = 0.$$

The function $f_\lambda(P)$ is said to converge in mean over all space to the function $f(P)$ as $\lambda \to \infty$.

We shall imply all of the above facts when we write:

$$f(P) = \underset{\lambda \to \infty}{\text{l.m.}} f_\lambda(P).$$

If $f(P) = \underset{\lambda \to \infty}{\text{l.m.}} f_\lambda(P)$ and if $g(P)$ is q.i. over ∞, Schwarz's inequality tells us that

$$\left| \int_{\infty} \{f(P) - f_\lambda(P)\} g(P) dV_P \right|^2 \leqq \int_{\infty} |f(P) - f_\lambda(P)|^2 dV_P \cdot \int_{\infty} |g(P)|^2 dV_P,$$

and hence that

(i) $$\lim_{\lambda \to \infty} \int_{\infty} f_\lambda(P) g(P) dV_P = \int_{\infty} f(P) g(P) dV_P.$$

In particular, setting $g(P)$ equal to unity when P lies inside a certain finite region V and equal to zero elsewhere, we find that

(ii) $$\lim_{\lambda \to \infty} \int_{V} f_\lambda(P) dV_P = \int_{V} f(P) dV_P.$$

Furthermore, since

$$|f(P)|^2 - |f_\lambda(P)|^2 = f(P)\{\bar{f}(P) - \bar{f}_\lambda(P)\} + \bar{f}(P)\{f(P) - f_\lambda(P)\} - |f(P) - f_\lambda(P)|^2,$$

where dashes denote conjugate values, we have:

(iii) $$\lim_{\lambda \to \infty} \int_{\infty} |f_\lambda(P)|^2 dV_P = \int_{\infty} |f(P)|^2 dV_P.$$

If, in addition, $g(P) = \underset{\lambda \to \infty}{\text{l.m.}} g_\lambda(P)$, then, since

$$f(P)g(P) - f_\lambda(P)g_\lambda(P) = \{f(P) - f_\lambda(P)\}g(P) + \{g(P) - g_\lambda(P)\}f_\lambda(P),$$

(iii) and Schwarz's inequality tell us that

$$\text{(iv)} \qquad \lim_{\lambda \to \infty} \int_{\infty} f_\lambda(P) g_\lambda(P) dV_P = \int_{\infty} f(P) g(P) dV_P.$$

Throughout the remainder of the paper we shall refer to this group of one theorem and four corollaries as the "theory of mean convergence."

First given as a special case of a more general transform theory by Plancherel,[6] the Fourier transform theorem was later established in simpler fashion by Titchmarsh.[7] The following is the complex, n-dimensional form of the

FOURIER TRANSFORM THEOREM.

(i) If $f(P)$ is q.i. over ∞, there exists its Fourier transform

$$F(Q) = \text{l.m.}_{V \to \infty} \frac{1}{(2\pi)^{n/2}} \int_V f(P) e^{i(P \cdot Q)} dV_P$$

a function defined uniquely almost everywhere, q.i. over ∞, and such that

$$\int_{\infty} |F(Q)|^2 dV_Q = \int_{\infty} |f(P)|^2 dV_P.$$

(ii) If $f(P)$ and $g(P)$ are each q.i. over ∞ and if their respective Fourier transforms are $F(Q)$ and $G(Q)$, then

$$\int_{\infty} F(Q) \overline{G}(Q) dV_Q = \int_{\infty} f(P) \overline{g}(P) dV_P.$$

(iii) If $f(P)$ is q.i. over ∞ and if $F(Q)$ is its Fourier transform, then the Fourier transform of $F(Q)$ is $f(-M)$, i.e.

$$f(M) = \text{l.m.}_{V \to \infty} \frac{1}{(2\pi)^{n/2}} \int_V F(Q) e^{-i(M \cdot Q)} dV_Q.$$

(iv) If $f(P)$ is q.i. over ∞, an explicit formula for its Fourier

[6] M. Plancherel: "Contributions à l'étude de la representation d'une fonction par des intégrales définies," Rendiconti del Circolo Matematico di Palermo, 30 (1910), p. 289.

[7] E. C. Titchmarsh: "Hankel transforms," Proc. Cambr. Philos. Soc., 21 (1923), p. 463; "A contribution to the theory of Fourier transforms," London Math. Soc. Proc. (2), 23 (1924), p. 279.

transform is

$$F(Q) = \lim_{A \to 0} \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}A^{n/2}} \int_\infty f(P) \frac{J_{n/2}(A|P|)}{|P|^{n/2}} e^{i(P\cdot Q)} dV_P,$$

this limit existing for almost every Q.

Proof of (i):

Construct the approximating function

$$f_V(P) = \begin{cases} f(P), \text{ when } P \text{ is in } V; \\ 0, \text{ when } P \text{ is not in } V. \end{cases}$$

From the integrability of $f(P)$ over any finite region there follows the existence of $F_V(Q)$, the Fourier transform of $f_V(P)$

$$F_V(Q) = \frac{1}{(2\pi)^{n/2}} \int_\infty f_V(P) e^{i(P\cdot Q)} dV_P.$$

It is important to note that this and other similar integrals to follow are infinite in appearance only; we may, at pleasure, invert the order of integration of iterated integrals of this type.

We shall now prove the important relation

$$\int_\infty |F_V(Q)|^2 dV_Q = \int_\infty |f_V(P)|^2 dV_P = \int_V |f(P)|^2 dV_P.$$

We have

$$\begin{aligned} |F_V(Q)|^2 &= F_V(Q) \cdot \bar{F}_V(Q) \\ &= \frac{1}{(2\pi)^n} \int_\infty f_V(P) e^{i(P\cdot Q)} dV_P \int_\infty \bar{f}_V(P') e^{-i(P'\cdot Q)} dV_{P'} \\ &= \frac{1}{(2\pi)^n} \int_\infty f_V(P) dV_P \int_\infty \bar{f}_V(P - M) e^{i(M\cdot Q)} dV_M \\ &= \frac{1}{(2\pi)^n} \int_\infty \left\{ e^{i(M\cdot Q)} \int_\infty f_V(P) \bar{f}_V(P - M) dV_P \right\} dV_M. \end{aligned}$$

Let $V_{\to M}$ denote the region obtained from a given region V by translating V, as a whole, through the vector M; and let $(V \cdot V_{\to M})$ denote the region common to V and to $V_{\to M}$. Then we see that the product $f_V(P) f_V(P - M)$ differs from zero only when P is in $(V \cdot V_{\to M})$.

Choose a radius B of such size that $(V \cdot V_{\rightarrow M})$ will be a null region for all vectors M of absolute value greater than B. Then

$$(V \cdot V_{\rightarrow M}) = 0 \quad \text{when} \quad |M - N| > 2B \text{ and } |N| \leqq B.$$

Hence

$$|F_V(Q)|^2 = \frac{1}{(2\pi)^n} \int_{|M-N| \leqq 2B} \left\{ e^{i(M \cdot Q)} \int_\infty f_V(P) \bar{f}_V(P - M) dV_P \right\} dV_M,$$
$$\text{when } |N| \leqq B.$$

Since this expression is constant for all vectors N of absolute value not greater than B, its value will not be altered if we average it with respect to N over the sphere $|N| \leqq B$. Therefore

$$|F_V(Q)|^2 = \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2} B^n (2\pi)^n} \int_{|N| \leqq B} dV_N \int_{|M-N| \leqq 2B} \times \left\{ e^{i(M \cdot Q)} \int_\infty f_V(P) \bar{f}_V(P - M) dV_P \right\} dV_M.$$

We wish next to expand the quantity within the brace into a series uniformly convergent for all M and Q. To do this we introduce the auxiliary function

$$f_\square(P) = \begin{cases} f(P), \text{ when } P \text{ is in } V; \\ 0, \text{ when } P \text{ is in } \square \text{ but not in } V; \end{cases}$$

defined throughout an n-dimensional "cube" $\square$, whose center is at the origin and whose sides are of length $8B$. We shall define $f_\square(P)$ outside $\square$ by requiring that $f_\square(P)$ be an n-tuply periodic function. We note that in each cube of periodicity there will be a region corresponding to V, and that none of these regions can be made to have a point in common with $\square$ by translation through a vector M whose absolute value is not greater than $3B$. This being so, we can write

$$|F_V(Q)|^2 = \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2} B^n (2\pi)^n} \int_{|N| \leqq B} dV_N \int_{|M-N| \leqq 2B} \times \left\{ e^{i(M \cdot Q)} \int_\square f_\square(P) \bar{f}_\square(P - M) dV_P \right\} dV_M.$$

Now, $f_\square(P)$ is q.i. over $\square$. Hence,[8] the Fourier series corresponding to $f_\square(P)$ converges in mean over $\square$ to $f_\square(P)$:

$$f_\square(P) = \text{l.m. over}_{\lambda\to\infty} \square \sum_{|K|<\lambda} a_K e^{i(K\cdot P)}$$

where the vectors K all have components in the directions of the edges of $\square$ which are integral multiples of $\pi/8B$. Since the functions $e^{i(K\cdot P)}$ are orthogonal in $\square$, we have, by the theory of mean convergence, that

$$\begin{aligned}\int_V |f(P)|^2 dV_P &= \int_\square |f_\square(P)|^2 dV_P \\ &= \lim_{\lambda\to\infty} \sum_{|K|<\lambda} \sum_{|L|<\lambda} a_K \bar{a}_L \int_\square e^{i[(K-L)\cdot P]} dV_P \\ &= (8B)^n \sum |a_K|^2,\end{aligned}$$

this last series being, therefore, absolutely convergent. In similar fashion we find that

$$\int_\square f_\square(P)\bar{f}_\square(P-M)dV_P = (8B)^n \sum |a_K|^2 e^{i(K\cdot M)}.$$

Hence

$$\left\{ e^{i(M\cdot Q)} \int_\square f_\square(P)\bar{f}_\square(P-M)dV_P \right\} = (8B)^n \sum |a_K|^2 e^{i[(Q+K)\cdot M]}.$$

This series is uniformly convergent for all M and all Q since it is dominated by the absolutely convergent series $(8B)^n \sum |a_K|^2$.

Thus, we find that

$$\begin{aligned}|F_V(Q)|^2 &= (8B)^n \sum |a_K|^2 \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}B^n(2\pi)^n} \\ &\quad \times \int_{|N|\le B} dV_N \int_{|M-N|\le 2B} e^{i[(Q+K)\cdot M]} dV_M \\ &= (8B)^n \sum |a_K|^2 \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}B^n(2\pi)^n}\end{aligned}$$

[8] By the n-dimensional form of the theorems given by E. W. Hobson: "Theory of functions of a real variable," 2d ed., 2, pp. 573–578.

$$\times \int_{|N|\leq B} e^{i[(Q+K)\cdot N]} dV_N \int_{|M'|\leq 2B} e^{i[(Q+K)\cdot M']} dV_{M'}$$

$$= (8B)^n \sum |a_K|^2 \frac{\Gamma\left(\frac{n}{2}+1\right)(2B)^{n/2}}{\pi^{n/2}B^{n/2}} \times \frac{J_{n/2}(B|Q+K|)J_{n/2}(2B|Q+K|)}{|Q+K|^n}.$$

and this series is uniformly convergent for all Q.

Relation (III), proven at the beginning of the paper, now yields us our desired result:

$$\int_\infty |F_V(Q)|^2 dV_Q = (8B)^n \sum |a_K|^2 \frac{\Gamma\left(\frac{n}{2}+1\right)(2B)^{n/2}}{\pi^{n/2}(B)^{n/2}} \times \int_\infty \frac{J_{n/2}(B|Q+K|)J_{n/2}(2B|Q+K|)}{|Q+K|^n} dV_Q$$

$$= (8B)^n \sum |a_K|^2$$

$$= \int_V |f(P)|^2 dV_P = \int_\infty |f_V(P)|^2 dV_P.$$

Consider, now, the function

$$f_{V_1}(P) - f_{V_2}(P)$$

where V_1 and V_2 are any two regions each of which we shall presently allow to expand indefinitely according to the restrictions previously laid down. The Fourier transform of this function is

$$F_{V_1}(Q) - F_{V_2}(Q)$$

where $F_{V_1}(Q)$ and $F_{V_2}(Q)$ are the respective Fourier transforms of $f_{V_1}(P)$ and $f_{V_2}(P)$. We see at once that the preceding analysis applies to these two functions. Hence

$$\int_\infty |F_{V_1}(Q) - F_{V_2}(Q)|^2 dV_Q = \int_\infty |f_{V_1}(P) - f_{V_2}(P)|^2 dV_P$$

$$= \int_{V_1-(V_1\cdot V_2)} |f(P)|^2 dV_P + \int_{V_2-(V_1\cdot V_2)} |f(P)|^2 dV_P$$

$$\leqq 2 \int_{P \text{ outside } (V_1\cdot V_2)} |f(P)|^2 dV_P.$$

But, since $f(P)$ is q.i. over ∞, this last expression converges to zero as V_1 and V_2 expand indefinitely. Thus

$$\lim_{\substack{V_1 \to \infty \\ V_2 \to \infty}} \int_\infty |F_{V_1}(Q) - F_{V_2}(Q)|^2 dV_Q = 0.$$

Therefore, by the theory of mean convergence, there exists

$$F(Q) = \underset{V \to \infty}{\text{l.m.}}\ F_V(Q) = \underset{V \to \infty}{\text{l.m.}} \frac{1}{(2\pi)^{n/2}} \int_V f(P) e^{i(P \cdot Q)} dV_P,$$

a function defined uniquely almost everywhere, q.i. over ∞, and such that

$$\int_\infty |F(Q)|^2 dV_Q = \lim_{V \to \infty} \int_\infty |F_V(Q)|^2 dV_Q$$

$$= \lim_{V \to \infty} \int_V |f(P)|^2 dV_P = \int_\infty |f(P)|^2 dV_P.$$

Proof of (*ii*):

Let $f(P)$ and $g(P)$ each be q.i. over ∞. Denote their Fourier transforms by $F(Q)$ and $G(Q)$, respectively. Then, if a and b are any two complex constants, the function

$$af(P) + bg(P)$$

will be q.i. over ∞, and will have for its Fourier transform, the function

$$aF(Q) + bG(Q).$$

Hence

$$\int_\infty |aF(Q) + bG(Q)|^2 dV_Q = \int_\infty |af(P) + bg(P)|^2 dV_P.$$

Now, since

$$f(P)\bar{g}(P) = \frac{1}{2}|f(P) + g(P)|^2 + \frac{i}{2}|f(P) + ig(P)|^2$$

$$- \frac{1+i}{2}|f(P)|^2 - \frac{1+i}{2}|g(P)|^2,$$

with a similar relation for $F(Q)\bar{G}(Q)$, it follows that

$$\int_\infty F(Q)\bar{G}(Q) dV_Q = \int_\infty f(P)\bar{g}(P) dV_P.$$

It is important to note the fact that the Fourier transform of $\bar{g}(P)$ is not $\bar{G}(Q)$ but $\bar{G}(-Q)$.

Proof of (iii):

Let $f(P)$ and $g(P)$ each be q.i. over ∞. Denote their Fourier transforms by $F(Q)$ and $G(Q)$, respectively. Denote the Fourier transforms of $F(Q)$ and $G(Q)$ by $\mathscr{F}(M)$ and $\mathscr{G}(M)$, respectively. Applying the second part of our theorem we see that

$$\int_{\infty} \mathscr{F}(M)\bar{\mathscr{G}}(M)dV_M = \int_{\infty} F(Q)\bar{G}(Q)dV_Q = \int_{\infty} f(P)\bar{g}(P)dV_P.$$

Now, consider the special function

$$g(P) = \begin{cases} 1, \text{ when } |P + M_0| < R; \\ 0, \text{ when } |P + M_0| > R. \end{cases}$$

It is q.i. over ∞, and its Fourier transform is

$$G(Q) = \frac{1}{(2\pi)^{n/2}} \int_{|P+M_0| \leq R} e^{i(P \cdot Q)} dV_P = \frac{R^{n/2} J_{n/2}(R|Q|)}{|Q|^{n/2}} e^{-i(M_0 \cdot Q)}$$

and (II) tells us that the Fourier transform of $G(Q)$ is

$$\begin{aligned} \mathscr{G}(M) &= \lim_{S \to \infty} \left(\frac{R}{2\pi}\right)^{n/2} \int_{|Q| \leq S} \frac{J_{n/2}(R|Q|)}{|Q|^{n/2}} e^{i[(M-M_0) \cdot Q]} dV_Q \\ &= \begin{cases} 1, \text{ when } |M - M_0| < R; \\ 0, \text{ when } |M - M_0| > R. \end{cases} \end{aligned}$$

For this $g(P)$ and its $\mathscr{G}(M)$ we find, then, that

$$\begin{aligned} \int_{|M-M_0| \leq R} \mathscr{F}(M)dV_m &= \int_{|P+M_0| \leq R} f(P)dV_P \\ &= \int_{|M-M_0| \leq R} f(-M)dV_m. \end{aligned}$$

Therefore

$$\int_{|M-M_0| \leq R} \{\mathscr{F}(M) - f(-M)\}dV_m = 0$$

for every radius R and every center M_0. In other words, the functions $\mathscr{F}(M)$ and $f(-M)$ can differ only on a set of zero measure, which is what we wished to prove

Proof of (iv):

Let $f(P)$ be q.i. over ∞ and let $F(Q)$ be its Fourier transform

$$F(M) = \underset{V\to\infty}{\text{l.m.}}\ F_V(M) = \underset{V\to\infty}{\text{l.m.}}\ \frac{1}{(2\pi)^{n/2}}\int_V f(P)e^{i(P\cdot M)}dV_P.$$

By the theory of mean convergence

$$\begin{aligned}\int_{|M-Q|\leq A} F(M)dV_m &= \lim_{V\to\infty}\int_{|M-Q|\leq A} F_V(M)dV_m \\ &= \lim_{V\to\infty}\frac{1}{(2\pi)^{n/2}}\int_V f(P)dV_P\int_{|M-Q|\leq A} e^{i(P\cdot M)}dV_m \\ &= A^{n/2}\int_\infty f(P)\frac{J_{n/2}(A|P|)}{|P|^{n/2}}e^{i(P\cdot Q)}dV_P,\end{aligned}$$

an absolutely convergent integral. Since $F(M)$ is integrable, it follows that the limit of the average of $F(M)$ over a region as the region shrinks around a point $M = Q$ is almost everywhere equal to $F(Q)$. Hence, almost everywhere

$$\begin{aligned}F(Q) &= \lim_{A\to 0}\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}A^n}\int_{|M-Q|\leq A} F(M)dV_m \\ &= \lim_{A\to 0}\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}A^{n/2}}\int_\infty f(P)\frac{J_{n/2}(A|P|)}{|P|^{n/2}}e^{i(P\cdot Q)}dV_P.\end{aligned}$$

Note.—To be strictly accurate in discussing the mean convergence of a volume integral as the region V of integration expands indefinitely one should require V to belong to a one-parameter family of regions. Thus more than one Fourier transform can be constructed. However, Schwarz's inequality shows at once that any two such transforms are equivalent, i. e. differ only on a set of zero measure.

A METHOD FOR DERIVING FORMULAS FOR THE APPROXIMATE CALCULATION OF INTEGRALS
(*Continued*) *

By Prescott D. Crout

1. **Euler's Formula and a Formula Analogous to That of Euler for the Approximate Calculation of Integrals.** In the first paper [1] the discussion of Euler's formula (1′) for the approximation of $\int_{-\theta}^{\theta} f(x)dx$ was carried to the point where a system of linear equations was obtained by which the coefficients could be determined.

This paper has two objects in view. The first is to continue the discussion of Euler's formula, finally obtaining the coefficients expressed in terms of the Bernoullian numbers. It will be shown that these coefficients are identical with the corresponding terms of the expansion of $\theta \coth \theta$ in powers of θ, or that they differ only in sign from the corresponding terms of the power series expansion of $\theta \cot \theta$.

The second object of this paper is to derive a certain formula (19′) for the approximation of $\int_{-\theta}^{\theta} f(x)dx$ which is analogous to that of Euler, the difference being that Euler's formula involves derivatives of order 0, 1, 3, 5, $\cdots$, $2k-3$ at the end points of the interval of integration, whereas this formula involves derivatives of order 0, 2, 4, 6, $\cdots$, $2k-2$ at the end points of that interval. It will be shown that the coefficients of (19′) may be expressed very simply in terms of the Bernoullian numbers; also that these coefficients are identical with the corresponding terms of the expansion of $\tanh \theta$ in powers of θ, or that they

* This paper was prepared under the supervision of Prof. George Rutledge, of the Massachusetts Institute of Technology, who offered many valuable suggestions.

[1] *Crout: A Method for Deriving Formulas for the Approximate Calculation of Integrals*, this *Journal*, vol. 7, p. 126.

differ only in sign from the corresponding terms of the power series expansion of tan θ.

Throughout this paper the methods and the notation will be that of the first paper, to which all section numbers and equation numbers refer. To avoid ambiguity, the equation numbers of this paper will be followed by a prime.

2. **Euler's Formula.** Euler's formula for k elements is of the form

$$\begin{aligned}\int_{-\theta}^{\theta} f(x)dx = {} & \theta[\{K_{-1}^{0}\}f(-\theta) + \{K_1^{0}\}f(\theta)] \\ & + \theta^2[\{K_{-1}^{1}\}f^1(-\theta) + \{K_1^{1}\}f^1(\theta)] \\ & + \theta^4[\{K_{-1}^{3}\}f^3(-\theta) + \{K_1^{3}\}f^3(\theta)] \\ & \cdots\cdots\cdots\cdots\cdots\cdots \\ & + \theta^{2k-2}[\{K_{-1}^{2k-3}\}f^{2k-3}(-\theta) + \{K_1^{2k-3}\}f^{2k-3}(\theta)] + \epsilon,\end{aligned} \tag{1'}$$

or

$$\begin{aligned}\int_{-\theta}^{\theta} f(x)dx = {} & \theta[\{K_{-1}^{0}\}f(-\theta) + \{K_1^{0}\}f(\theta)] \\ & + \sum_{i=2}^{k} \theta^{2i-2}[\{K_{-1}^{2i-3}\}f^{2i-3}(-\theta) + \{K_1^{2i-3}\}f^{2i-3}(\theta)] \\ & + \epsilon\end{aligned}$$

where $k > 1$. From Section 9 we have

$$\begin{aligned}\{K_{-1}^{0}\} &= \{K_1^{0}\} = 1 \\ -\{K_{-1}^{2i-3}\} &= \{K_1^{2i-3}\} \quad i = 2, 3, 4, \cdots, k,\end{aligned} \tag{2'}$$

and

$$\begin{aligned}1 &= \{K_1^{0}\}\{1, 0\} \\ 1 &= \{K_1^{0}\}\{3, 0\} + \{K_1^{1}\}\{3, 1\} \\ 1 &= \{K_1^{0}\}\{5, 0\} + \{K_1^{1}\}\{5, 1\} + \{K_1^{3}\}\{5, 3\} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 1 &= \{K_1^{0}\}\{2k-1, 0\} + \{K_1^{1}\}\{2k-1, 1\} \\ &\quad + \{K_1^{3}\}\{2k-1, 3\} + \cdots + \{K_1^{2k-3}\}\{2k-1, 2k-3\}.\end{aligned} \tag{3'}$$

The determinant of the first $j+1$ equations of (3') is the product of the elements of the principal diagonal, or

$$\Delta_{j+1} = \prod_{i=1}^{j} \overline{2i+1}! \neq 0 \quad j = 1, 2, 3, \cdots, k-1;$$

hence we can solve these $j+1$ equations for $\{K_1^{2i-1}\}$, $j>0$, and thus have

$$\{K_1^{2i-1}\} = \frac{1}{\prod_{i=1}^{j} \overline{2i+1}!} \times \begin{vmatrix} \{1,0\} & 0 & 0 & \cdots & 0 & 1 \\ \{3,0\} & \{3,1\} & 0 & \cdots & 0 & 1 \\ \{5,0\} & \{5,1\} & \{5,3\} & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \{2j-1,0\} & \{2j-1,1\} & \{2j-1,3\} & \cdots & \{2j-1,2j-3\} & 1 \\ \{2j+1,0\} & \{2j+1,1\} & \{2j+1,3\} & \cdots & \{2j+1,2j-3\} & 1 \end{vmatrix}. \quad (4')$$

Noting that $\{n, m\} = \overline{m+1}!\binom{n}{m+1}$, we have from (4')

$$\{K_1^{2i-1}\} = \frac{\prod_{i=1}^{j-1} \overline{2i}!}{\prod_{i=1}^{j} \overline{2i+1}!} \times \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 3 & 3 & 0 & \cdots & 0 & 1 \\ 5 & \binom{5}{2} & 5 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 2j-1 & \binom{2j-1}{2} & \binom{2j-1}{4} & \cdots & 2j-1 & 1 \\ 2j+1 & \binom{2j+1}{2} & \binom{2j+1}{4} & \cdots & \binom{2j+1}{2j-2} & 1 \end{vmatrix}, \quad (5')$$

or

$$\{K_1^{2i-1}\} = \frac{D_j}{\overline{2j}! \prod_{i=1}^{j} (2i+1)} \qquad j = 1, 2, 3, \cdots, k-1 \quad (6')$$

where D_j denotes the determinant in (5').

We now consider the recursion formula [2]

$$2h = \sum_{i=1}^{h} 2^{2i} \binom{2h+1}{2i} B_{2i} \qquad h = 1, 2, 3, \cdots \quad (7')$$

[2] *Saalschütz: Vorlesungen über die Bernoullischen Zahlen*, p. 12, formula V.

where B_2, B_4, B_6, $\cdots$ are the Bernoullian numbers (13′); and write (7′) in the form

$$\begin{aligned} -1 &= A \\ -1 &= (2h+1)A + \sum_{i=1}^{h} 2^{2i} \binom{2h+1}{2i} B_{2i} \quad h = 1, 2, 3, \cdots. \end{aligned} \tag{8′}$$

Since the determinant of the first $j+1$ equations of (8′) is the product of the elements of the principal diagonal, or

$$\Delta_{j+1} = \prod_{i=1}^{j} 2^{2i}(2i+1) \neq 0 \quad j = 1, 2, 3, \cdots,$$

we can determine B_{2j} from these $j+1$ equations; thus

$$B_{2j} = \frac{-1}{2^{2j} \prod_{i=1}^{j} (2i+1)} \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 3 & 3 & 0 & \cdots & 0 & 1 \\ 5 & \binom{5}{2} & 5 & \cdots & 0 & 1 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 2j-1 & \binom{2j-1}{2} & \binom{2j-1}{4} & \cdots & 2j-1 & 1 \\ 2j+1 & \binom{2j+1}{2} & \binom{2j+1}{4} & \cdots & \binom{2j+1}{2j-2} & 1 \end{vmatrix}, \tag{9′}$$

or

$$D_j = -B_{2j} 2^{2j} \prod_{i=1}^{j} (2i+1) \quad j = 1, 2, 3, \cdots. \tag{10′}$$

Substituting (10′) in (6′), noting (2′), we have

$$-\{K_{-1}^{2j-1}\} = \{K_1^{2j-1}\} = \frac{-2^{2j} B_{2j}}{2j!} \quad j = 1, 2, 3, \cdots, k-1; \tag{11′}$$

thus finally [3],

$$\begin{aligned} \int_{-\theta}^{\theta} f(x)dx = \theta[f(-\theta) + f(\theta)] &+ \sum_{j=1}^{k-1} \frac{2^{2j} B_{2j} \theta^{2j}}{2j!} [f^{2j-1}(-\theta) - f^{2j-1}(\theta)] \\ &+ \epsilon, \end{aligned} \tag{12′}$$

[3] This expression was to be expected due to Section 12, *d*, since (12′) is a familiar form of Euler's formula. *La Vallée Poussin, Cours d'Analyse Infinitésimale*, vol. 2, p. 343. *Jordan, Cours d'Analyse*, vol. 2, p. 119.

due to (11′) and (1′), where the Bernoullian numbers B_{2j} are

$$B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30},$$
$$B_{10} = \frac{5}{66}, \quad B_{12} = -\frac{691}{2730}, \quad B_{14} = \frac{7}{6}, \quad B_{16} = -\frac{3617}{510}, \quad \cdots. \tag{13′}$$

Since

$$\theta \coth \theta = 1 + \sum_{j=1}^{\infty} \frac{2^{2j} B_{2j} \theta^{2j}}{\overline{2j}!}, \tag{14′}$$

it is evident that *the coefficient of* $[f^{2j-1}(-\theta) - f^{2j-1}(\theta)]$ *in* (12′) *is the* $j+1^{\text{th}}$ *term of the expansion* [4] *of* $\theta \coth \theta$ *in powers of* θ; similarly since

$$\theta \cot \theta = 1 + \sum_{j=1}^{\infty} \frac{(-1)^j 2^{2j} B_{2j} \theta^{2j}}{\overline{2j}!}, \tag{15′}$$

it is evident that *the coefficient of* $[f^{2j-1}(-\theta) - f^{2j-1}(\theta)]$ *in* (12′) *is* $(-1)^j$ *times the* $j+1^{\text{th}}$ *term of the expansion* [4] *of* $\theta \cot \theta$ *in powers of* θ; *thus we have relations involving the* $j+1^{\text{th}}$ *terms of* (12′), (14′), *and* (15′). For example, since

$$\begin{aligned} \theta \coth \theta &= 1 + \frac{\theta^2}{3} - \frac{\theta^4}{45} + \frac{2\theta^6}{945} - \frac{\theta^8}{4725} + \cdots, \\ \theta \cot \theta &= 1 - \frac{\theta^2}{3} - \frac{\theta^4}{45} - \frac{2\theta^6}{945} - \frac{\theta^8}{4725} - \cdots, \end{aligned} \tag{16′}$$

(12′) may be written

$$\begin{aligned} \int_{-\theta}^{\theta} f(x)dx = {} & \theta[f(-\theta) + f(\theta)] + \frac{\theta^2}{3}[f^1(-\theta) - f^1(\theta)] \\ & - \frac{\theta^4}{45}[f^3(-\theta) - f^3(\theta)] + \frac{2\theta^6}{945}[f^5(-\theta) - f^5(\theta)] \\ & - \frac{\theta^8}{4725}[f^7(-\theta) - f^7(\theta)] + \cdots \\ & + \frac{2^{2k-2} B_{2k-2} \theta^{2k-2}}{\overline{2k-2}!}[f^{2k-3}(-\theta) - f^{2k-3}(\theta)] + \epsilon. \end{aligned} \tag{17′}$$

[4] The series (14′) and (15′) have the same region of convergence, $-\pi < \theta < \pi$; however, it is evident that these series are used here in such a way that their convergence is of no importance.

From (84) we see that $\Omega = 2k$; thus (12′), *or* (17′), *gives exact results when applied to any polynomial of degree* $2k - 1$.

In using a formula of approximate integration, it is obvious from (68) that ϵ can be reduced by breaking the interval of integration up into subintervals and by then applying the formula to each partial integral, the results being finally added; however Ω remains unaltered. Simpson's rule is an example of this process; as another example let us approximate $\int_0^{2n\theta} f(x)dx$ by applying (12′), or (17′), to each of the partial integrals formed by breaking the interval of integration $(0, 2n\theta)$ up into n equal subintervals. We then have

$$\int_0^{2n\theta} f(x)dx = \theta\left[f(0) + 2\sum_{i=1}^{n-1} f(2i\theta) + f(2n\theta)\right] + \sum_{j=1}^{k-1} \frac{2^{2j}B_{2j}\theta^{2j}}{2j!}[f^{2j-1}(0) - f^{2j-1}(2n\theta)] + \epsilon_1. \tag{18′}$$

The substitution $2\theta = h$ places (18′) in the usual form of Euler's summation formula.[5]

3. **Formula Analogous to That of Euler.** The formula of integration which we shall consider is of the form

$$\begin{aligned}\int_{-\theta}^{\theta} f(x)dx = {} & \theta[\{K_{-1}{}^0\}f(-\theta) + \{K_1{}^0\}f(\theta)] \\ & + \theta^3[\{K_{-1}{}^2\}f^2(-\theta) + \{K_1{}^2\}f^2(\theta)] \\ & \cdots\cdots\cdots\cdots \\ & + \theta^{2k-1}[\{K_{-1}{}^{2k-2}\}f^{2k-2}(-\theta) + \{K_1{}^{2k-2}\}f^{2k-2}(\theta)] \\ & + \epsilon,\end{aligned} \tag{19′}$$

or

$$\int_{-\theta}^{\theta} f(x)dx = \sum_{i=1}^{k} \theta^{2i-1}[\{K_{-1}{}^{2i-2}\}f^{2i-2}(-\theta) + \{K_1{}^{2i-2}\}f^{2i-2}(\theta)] + \epsilon;$$

[5] *La Vallée Poussin: Cours d'Analyse Infinitésimale*, vol. 2, p. 344. *Pierpont: Functions of a Complex Variable*, p. 318.

hence, applying the theory developed in Section 8, we have

$$S = \sum_{i=1}^{k} \theta^{2i-1}[\{K_{-1}^{2i-2}\}f^{2i-2}(-\theta) + \{K_1^{2i-2}\}f^{2i-2}(\theta)], \quad (20')$$

from which it follows, due to (65), that

$$\{K_{-1}^{2i-2}\} = \{K_1^{2i-2}\} \quad i = 1, 2, 3, \cdots, k. \quad (21')$$

The following system is due to (60), (56), and (57),

$$\begin{aligned}
1 &= \{K_1^0\}\{1, 0\} + \{K_1^2\}\{1, 2\} + \cdots + \{K_1^{2k-2}\}\{1, 2k-2\} \\
1 &= \{K_1^0\}\{3, 0\} + \{K_1^2\}\{3, 2\} + \cdots + \{K_1^{2k-2}\}\{3, 2k-2\} \\
1 &= \{K_1^0\}\{5, 0\} + \{K_1^2\}\{5, 2\} + \cdots + \{K_1^{2k-2}\}\{5, 2k-2\} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
1 &= \{K_1^0\}\{2k-1, 0\} + \{K_1^2\}\{2k-1, 2\} + \cdots \\
&\qquad + \{K_1^{2k-2}\}\{2k-1, 2k-2\}.
\end{aligned} \quad (22')$$

Noting that $\{n, m\} = 0$ if $n < m + 1$, we see that all terms of (22′) which lie above the principal diagonal are zero; hence the determinant of the first j equations of (22′) is the product of the elements of the principal diagonal, or

$$\Delta_j = \prod_{i=1}^{j} \overline{2i-1}! \neq 0 \quad j = 1, 2, 3, \cdots, k. \quad (23')$$

It is thus evident that $\{K_1^{2j-2}\}$ may be determined from the first j equations of (22′); then, using determinants, we have

$$\{K_1^{2j-2}\} = \frac{1}{\prod_{i=1}^{j} \overline{2i-1}!} \times \begin{vmatrix} \{1, 0\} & 0 & \cdots & 0 & 1 \\ \{3, 0\} & \{3, 2\} & \cdots & 0 & 1 \\ \{5, 0\} & \{5, 2\} & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \{2j-3, 0\} & \{2j-3, 2\} & \cdots & \{2j-3, 2j-4\} & 1 \\ \{2j-1, 0\} & \{2j-1, 2\} & \cdots & \{2j-1, 2j-4\} & 1 \end{vmatrix}. \quad (24')$$

Noting that $\{n, m\} = \overline{m+1}!\binom{n}{m+1}$, we have from (24′)

$$\{K_1^{2j-2}\} = \frac{1}{\overline{2j-1}!}\begin{vmatrix} 1 & 0 & \cdots & 0 & 1 \\ \binom{3}{1} & 1 & \cdots & 0 & 1 \\ \binom{5}{1} & \binom{5}{3} & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \binom{2j-3}{1} & \binom{2j-3}{3} & \cdots & 1 & 1 \\ \binom{2j-1}{1} & \binom{2j-1}{3} & \cdots & \binom{2j-1}{2j-3} & 1 \end{vmatrix}; \quad (25')$$

also

$$\beta_{2j} = (-1)^{j-1}\begin{vmatrix} 1 & 0 & \cdots & 0 & 1 \\ \binom{3}{1} & 1 & \cdots & 0 & 1 \\ \binom{5}{1} & \binom{5}{3} & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \binom{2j-3}{1} & \binom{2j-3}{3} & \cdots & 1 & 1 \\ \binom{2j-1}{1} & \binom{2j-1}{3} & \cdots & \binom{2j-1}{2j-3} & 1 \end{vmatrix} \quad (26')$$

where

$$\tan x = \sum_{j=1}^{\infty} \beta_{2j} \frac{x^{2j-1}}{\overline{2j-1}!} \qquad -\frac{\pi}{2} < x < \frac{\pi}{2},$$

and

$$B_{2j} = \frac{(-1)^{j-1}2j}{2^{2j}(2^{2j}-1)}\beta_{2j} \quad (27')$$

where B_2, B_4, B_6, $\cdots$ are the Bernoullian numbers (13′).[6] Substituting (26′) and (27′) in (25′), noting (21′), we have

$$\{K_{-1}^{2j-2}\} = \{K_1^{2j-2}\} = \frac{2^{2j}(2^{2j}-1)}{\overline{2j}!}B_{2j} \quad j = 1, 2, 3, \cdots, k; \quad (28')$$

[6] *Pascal: Die Determinanten*, p. 136.

thus from (28′) and (19′) we have finally

$$\int_{-\theta}^{\theta} f(x)dx = \sum_{j=1}^{k} \frac{2^{2j}(2^{2j}-1)B_{2j}}{\overline{2j}!}\theta^{2j-1}[f^{2j-2}(-\theta)+f^{2j-2}(\theta)] + \epsilon, \tag{29′}$$

where the Bernoullian numbers B_{2j} are given by (13′). Since

$$\tanh\theta = \sum_{j=1}^{\infty} \frac{2^{2j}(2^{2j}-1)B_{2j}}{\overline{2j}!}\theta^{2j-1}, \tag{30′}$$

it is evident that *the coefficient of* $[f^{2j-2}(-\theta)+f^{2j-2}(\theta)]$ *in* (29′) *is the* j^{th} *term of the expansion*[7] *of* $\tanh\theta$ *in powers of* θ; similarly, since

$$\tan\theta = \sum_{j=1}^{\infty} \frac{(-1)^{j-1}2^{2j}(2^{2j}-1)B_{2j}}{\overline{2j}!}\theta^{2j-1}, \tag{31′}$$

it is evident that *the coefficient of* $[f^{2j-2}(-\theta)+f^{2j-2}(\theta)]$ *in* (29′) *is* $(-1)^{j-1}$ *times the* j^{th} *term of the expansion*[7] *of* $\tan\theta$ *in powers of* θ. For example, since

$$\begin{aligned}\tanh\theta &= \theta - \frac{\theta^3}{3} + \frac{2\theta^5}{15} - \frac{17\theta^7}{315} + \frac{62\theta^9}{2835} - \cdots \\ \tan\theta &= \theta + \frac{\theta^3}{3} + \frac{2\theta^5}{15} + \frac{17\theta^7}{315} + \frac{62\theta^9}{2835} + \cdots,\end{aligned} \tag{32′}$$

(29′) may be written

$$\begin{aligned}\int_{-\theta}^{\theta} f(x)dx = {} & \theta[f(-\theta)+f(\theta)] - \frac{\theta^3}{3}[f^2(-\theta)+f^2(\theta)] \\ & + \frac{2\theta^5}{15}[f^4(-\theta)+f^4(\theta)] - \frac{17\theta^7}{315}[f^6(-\theta)+f^6(\theta)] \\ & + \frac{62\theta^9}{2835}[f^8(-\theta)+f^8(\theta)] - \cdots \\ & + \frac{2^{2k}(2^{2k}-1)B_{2k}}{\overline{2k}!}\theta^{2k-1}[f^{2k-2}(-\theta)+f^{2k-2}(\theta)]+\epsilon.\end{aligned} \tag{33′}$$

[7] The series (30′) and (31′) have the same region of convergence, $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$; however, it is evident that these series are used here in such a way that their convergence is of no importance.

Also, we have

$$\begin{aligned} I &= 2k - 2 \\ T &= k \qquad \text{for (69)} \\ T &= k + 1 \quad \text{for (70)} \\ \Sigma_1 &= 0 \\ \Omega = \Sigma_2 &= 2k; \end{aligned} \tag{34'}$$

thus (29′), *or* (33′), *gives exact results when applied to any polynomial of degree* $2k - 1$.

EQUIMOMENTAL SYSTEMS[1]

By Philip Franklin

1. **Introduction.** The object of this note is to discuss certain simple geometric relations between rigid systems of mass-points which are equimomental, i.e. have the same moment of inertia about every axis in space.

While some of these results are in the literature,[2] they are relatively so little known that their interest justifies a connected account from an elementary standpoint.[3] Our introduction of the elementary equimomental operation of "sliding," and deductions therefrom, are believed to be new.

2. **Fundamental Properties of Moments of Inertia.** *Related ellipsoids.* For future reference, we recall some of the well-known elementary properties of moments of inertia.[4] From the theorem on parallel axes, and the dependence of the moments of inertia for one point on the coefficients for that point, there follows: *Two systems are equimomental if and only if they have the same total mass, center of gravity, and coefficients of inertia for the center of gravity.*

When the coördinate planes are taken to be the principal planes for the center of gravity, we have

$$(1) \qquad \Sigma mx = 0, \quad \Sigma my = 0, \quad \Sigma mz = 0; \qquad \Sigma myz = 0, \quad \Sigma mzx = 0, \quad \Sigma mxy = 0,$$

and the only essential coefficients of inertia are

$$(2) \quad A = \Sigma m(y^2 + z^2), \quad B = \Sigma m(z^2 + x^2), \quad C = \Sigma m(x^2 + y^2).$$

From these quantities, we may build up various ellipsoids, each of which determines the inertial properties of the system. In

[1] Presented to the American Mathematical Society, December, 1928.

[2] See, in particular, Reye, *Crelle*, vol. 72 (1870), p. 302 ff.

[3] Perhaps the most complete accessible account is E. J. Routh, "Rigid Dynamics," London, 1877, vol. I, Chap. 1, and he confines himself to special cases of our theorems.

[4] Cf. e.g. Routh, l.c. or Appell, "Mecanique Rationelle," vol. 2, chap. 1.

the first place, we have the ellipsoid of inertia,

$$Ax^2 + By^2 + Cz^2 = 1, \tag{3}$$

for which if r is the radius vector in any direction, $1/r^2$ is the moment of inertia about the corresponding axis. Secondly we have the ellipsoid of gyration,

$$\frac{x^2}{A} + \frac{y^2}{B} + \frac{z^2}{C} = 1, \tag{4}$$

for which, if the perpendicular on the tangent plane in any direction is p, the moment of inertia about the corresponding perpendicular is p^2. Geometrically, (3) and (4) are polar reciprocals. Thirdly we have the less known class of ellipsoids,

$$\frac{x^2}{B + C - A} + \frac{y^2}{C + A - B} + \frac{z^2}{A + B - C} = \frac{5}{2kM}. \tag{5}$$

If $k = 1$, a solid, homogeneous ellipsoid with this equation and mass M, the total mass of the system, is equimomental to the system, so that we may call it the equivalent ellipsoid. If $k = 5/M$, the ellipsoid has the property that the perpendicular on any tangent plane is equal to the "moment of inertia" for the parallel plane through the center, i.e. the sum of each mass times the square of the perpendicular from it to the latter plane. Geometrically, the relation of (5) to (3) is given by the fact that the locus of points from which three mutually perpendicular tangent lines can be drawn to (5) is an ellipsoid [6] similar to (3), equal to it if:

$$4kM/5 = 2BC + 2CA + 2AB - A^2 - B^2 - C^2. \tag{6}$$

If $k = -5$, the imaginary ellipsoid (5) is confocal to

$$\frac{x^2}{A} + \frac{y^2}{B} + \frac{z^2}{C} = \frac{1}{M}, \tag{7}$$

similar to (4). The ellipsoids confocal to (7) may be characterized as the envelopes of planes for which the "moment of inertia"

[6] Cf. G. Salmon, "Analytic Geometry of three dimensions," 5th edition, revised by A. P. Rogers (1912), vol. 1, p. 177; or *American Mathematical Monthly*, vol. 34 (1927), p. 467.

of the system is constant. The lines through which two perpendicular tangent planes can be drawn to one of these confocals are a set of lines for which the moment of inertia of the system is constant.

3. **Elementary Equimomental Operations.** Let us consider a finite number of mass-points, of various masses, rigidly connected. What operations may we perform on the points, in the sense of replacing certain points by certain other points, with different masses, so as to leave the new system equimomental to the old? For a single point, we may evidently not do anything, except regard it as several coincident points, since no set of more than one point has all the moments of inertia for axes through the center of gravity zero.

For two points, P_1 with mass m_1 and P_2 with mass m_2 distant, respectively, km_2 and km_1 from their center of gravity, we may perform an operation of "sliding," i.e. replace them by a second pair of points on the same straight line, equimomental to the first pair. If the new points are P_1' with mass m_1' and P_2' with mass m_2', distant respectively $k'm_2'$ and $k'm_1'$ from the same center of gravity, we evidently must have

$$(8) \qquad m_1 + m_2 = m_1' + m_2'; \qquad k^2 m_1 m_2 = k'^2 m_1' m_2'.$$

It should be noted that these restrictions leave but one degree of freedom. That is, we may select arbitrarily one mass, m_1',

$$0 < m_1' < m_1 + m_2,$$

or one distance, $k'm_2' > 0$, or the ratio of the masses (or distances) $m_1'/m_2' > 0$; and the remaining primed constants will be uniquely determined from (8), with the new masses positive. In particular, one of the new particles may be given any position on the line except the center of gravity of the pair.

Now consider four mass points in a plane, A, B, C, D. If AB is parallel to CD, or meets it in the center of gravity of the pair C, D, a slight sliding of the pair B, D will disturb these relations. Hence we may assume that AB meets CD in E, not the center of gravity of the latter pair. We slide the pair C, D to a new pair E, F. Of the three A, B, E, at least one of the points A, B, say A, will not be between the other two, and hence

not the center of gravity of them, so that we may slide B, E to a new pair A', G, the first of which coincides in position with A. We may now coalesce the two masses at A into a single mass at that point, and we have reduced the system to three points. If we start with more than four particles in a plane, we may repeat the process just described with sets of four until only three points remain, and we have:

THEOREM I. *For any finite set of mass-points in a plane an equimomental set of three points may be found by slidings.*

COROLLARY. *One of the three points may be given an arbitrary position in the plane, not the center of gravity of the given set (assumed to be not all on a single line).*

For, let ABC be the triplet we first arrive at, and D the new position. Since D is not at the center of gravity of A, B, C, there is at least one median, say that from A, on which it does not lie. If this median cuts BC in E, by sliding B, C properly, one of these points may be brought so close to E that D is outside of the triangle ABC. By suitably lettering the triangle, and possibly a slight sliding of B, C to avoid parallelism, we may arrange matters so that DB cuts AC in F outside of AC, with D outside of BF. Consequently, we may slide A, C to F, G, and then B, F to D, H, so that the triplet A, B, C are replaced by D, G, H.

If two plane systems of three points (not all on a single line) are equimomental and have one point A of the first set in the same position as one point A' of the second set, the corresponding masses must agree, as well as the line of the remaining pairs, B, C and B', C'. For, suppose the mass at A, m_a less than that at A', m_a'. If h is the perpendicular distance from A (or A') to BC, the moment of inertia for the axis BC is $m_a h^2$ for the first system, while it is at least $m_a' h^2$ for the second. Thus an inequality in the masses at A, A' contradicts the equimomentality. Furthermore, if B' or C' were not on the line BC, the moment of inertia about this axis would be greater for the second system.

If two finite, plane sets of mass-points are given, not all on a line, we may reduce the first set to three, A, B, C and then using the corollary, the second set to three A', D, E such that the position of A' is the same as that of A. It follows then that the

masses are the same, and the line of D, E coincides with that of B, C. These pairs are clearly equimomental, and may be brought to coincidence by slidings, so that both sets may be reduced to the same triplet. If the slidings and combinations for one set are reversed, we get a series of slidings taking us from the first to the second set. If all the points of one set are on a line, those of any equimomental set must be on the same line and they may be reduced to the same pair by slidings. This proves:

THEOREM II. *A series of slidings and combinations may be found for any given finite, plane set of mass-points, which will take it into any other equimomental set.*

Now consider five mass-points in space, A, B, C, D, E. If the first four lie in a plane, we apply theorem I to them. If not, we may assume, or get by sliding B, D, that ED meets plane A, B, C in a point F not the center of gravity of A, B, C. By the corollary to theorem I, we may reduce A, B, C to a triplet of which F is one point, and by sliding reduce the three points D, E, F to two. Thus our five points are reduced to four. As the process may be repeated if the set we start with contains more than five points, we have:

THEOREM III. *For any finite set of mass-points in space an equimomental set of four points may be found by slidings.*

By reasoning analogous to that for two dimensions, we also have:

COROLLARY. *One of the four points may be given an arbitrary position in space, not the center of gravity of the given set (assumed to be not all in a single plane).*

The geometrical position of one of the four points determines the mass at that point as well as the plane of the other three. We prove this as for the plane case, using the "moment of inertia" for the plane of the other three. Evidently systems equimomental for all axes, are also equimomental for all planes. This leads to

THEOREM IV. *A series of slidings and combinations may be found for any given finite set of mass-points in space, which will take it into any other equimomental set.*

The dependence of mass on position for the minimum equi-

momental set, i.e. with the minimum number of points, gives the number of disposable parameters. The position of the first point in space gives three, the second in a plane two, and the third on a line one. This proves

THEOREM V. *In the construction of a minimum equimomental set, the number of arbitrary parameters is* 6, 3, 1 *or* 0, $[n(n-1)/2]$ *according as the number of points in the set is* 4, 3, 2 *or* 1, $[n]$.

We note that any discrete or continuous distribution is equimomental to its equivalent ellipsoid, and this in turn is equimomental to a finite set of mass-points, e.g. to six masses each $M/6$ located respectively at the points:

$$\left(\pm\sqrt{\frac{3(B+C-A)}{2M}},\ 0,\ 0\right);\quad \left(0,\ \pm\sqrt{\frac{3(C+A-B)}{2M}},\ 0\right);$$

$$\left(0,\ 0,\ \pm\sqrt{\frac{3(A+B-C)}{2M}}\right).$$

Thus theorems I, III and their corollaries are applicable to any distributions of masses.

Instead of considering the parameters for a minimum set as determined by the position of the points, we may regard them as coming, in part, from their masses. The mass of one point may be made small, by sliding it far from the center of gravity; and as large as any quantity less than the total mass by sliding the other three points far from the center of gravity. It follows that at least one minimum set may be found with any desired ratio of the masses. To investigate what happens when the masses are fixed, but the positions vary subject to this condition we turn to the appropriate transformation.

4. **Affine Transformations.** Temporarily using x_1, x_2, x_3 for x, y, z, we may write as the equations for an affine point transformation:

$$x_i' = \sum_j a_{ij}x_j + b_i \quad i, j = 1, 2, 3. \tag{9}$$

where the determinant

$$|a_{ij}| \neq 0. \tag{10}$$

We add an equation of transformation for densities:

$$\rho' = K\rho/|a_{ij}|, \tag{11}$$

so that

$$dm' = Kdm, \tag{12}$$

or $m' = Km$ for mass-points. On applying this transformation to a distribution of masses with total mass $M = \int dm$, coordinates of center of gravity $\bar{x}_i$, and products of inertia $A_{ij} = \int x_i x_j dm$, we find:

$$M' = KM, \qquad \bar{x}_i' = \sum_j a_{ij}\bar{x}_j + b_i, \qquad \text{and}$$

$$\bar{A}_{ij}' = K\Big[\sum_{p,q} a_{ip}a_{jq}A_{pq} + b_jM\sum_p a_{ip}\bar{x}_p + b_iM\sum_q a_{jq}\bar{x}_q + b_ib_jM\Big]. \tag{13}$$

These relations show that, if two systems have the same total mass, center of gravity, and products of inertia, their transforms will also have these quantities equal and we may state:

THEOREM VI. *The affine transforms of equimomental systems are equimomental.*

Since any system has a uniquely determined equivalent ellipsoid to which it is equimomental, it follows that

THEOREM VII. *An affine transformation takes a given system into an equimomental one if and only if it leaves the corresponding equivalent ellipsoid invariant.*

The affine transformation:

$$U\begin{cases} x' = x\sqrt{\dfrac{2M}{5(B+C-A)}}, \\ y' = y\sqrt{\dfrac{2M}{5(C+A-B)}}, \\ z' = z\sqrt{\dfrac{2M}{5(A+B-C)}}, \end{cases} \tag{14}$$

is of particular interest since it takes the equivalent ellipsoid (5) into the unit sphere. The affine transformations which leave this " ellipsoid " invariant are the rotations, R, about the center of gravity. Evidently the general transformation of theorem VII may be written:

$$T = U^{-1}\cdot R\cdot U, \tag{15}$$

so that, like the rotations, it depends on three parameters. These three parameters, together with the ratios of the four masses, make up the total of six for a minimum set. We obtain a canonical form for a minimum set with given masses m_1, m_2, m_3, m_4, with sum M, by using two of the additional parameters to put m_1 on the z-axis, and the third to put m_2 in the xz-plane. These conditions, together with:

$$\begin{aligned} &\sum_i m_i x_i = \sum_i m_i y_i = \sum_i m_i z_i = 0; \quad (i = 1, 2, 3, 4) \\ &\sum_i m_i y_i z_i = \sum_i m_i z_i x_i = \sum_i m_i x_i y_i = 0; \end{aligned} \tag{16}$$

yield the following coordinates:

$$\begin{array}{cccc} & x & y & z \\ m_1 & 0 & 0 & (m_2 + m_3 + m_4)s \\ m_2 & 0 & (m_3 + m_4)t & -m_1 s \\ m_3 & um_4 & -m_2 t & -m_1 s \\ m_4 & -um_3 & -m_2 t & -m_1 s \end{array} \tag{17}$$

A suitable choice of the constants s, t, u will give any required equivalent ellipsoid, since from (17):

$$\begin{aligned} &\sum_i m_i x_i^2 = u^2 m_3 m_4 (m_3 + m_4), \\ &\sum_i m_i y_i^2 = t^2 m_2 (m_3 + m_4)(m_2 + m_3 + m_4), \\ &\sum_i m_i z_i^2 = s^2 m_1 (m_2 + m_3 + m_4)(m_1 + m_2 + m_3 + m_4). \end{aligned} \tag{18}$$

If the equivalent ellipsoid is a sphere, these three sums are equal, and we find from (17) and (18) that the opposite edges of the tetrahedron in question are perpendicular, and the altitudes intersect in the center of gravity.[1] This property is not disturbed by a rotation, R, and since an affine transformation U^{-1} takes perpendicularity into conjugacy with respect to the transforms of spheres, i.e. the equivalent ellipsoid, we have:

THEOREM VIII. *If, through each vertex of the tetrahedron formed by the four mass-points of a minimum set, lines be drawn*

[1] For the properties of such orthogonal tetrahedra, cf. e.g. J. L. Coolidge, "A Treatise on the Circle and the Sphere," Oxford, 1916, p. 234; or *American Mathematical Monthly*, vol. 26 (1919), p.149.

conjugate, with respect to the equivalent ellipsoid, to their opposite faces, these lines will intersect, and at the center of gravity of the four mass-points.

COROLLARY. *When the ellipsoid is a sphere, the tetrahedron is orthogonal, and its ortho-center is the center of gravity of the mass-points*

For the plane case, these theorems have analogues about the location of the orthocenter, or intersection of conjugate lines at the center of gravity, but place no restriction on the shape of the triangle, since every triangle is " orthogonal " with respect to a circle, or any ellipse.

A special case of the corollary is that in which the masses are equal, and the tetrahedron is regular.

By considering the spherical case, or directly, we see that the ratio of the masses fixes the minimum set to within an affine transformation leaving the equivalent ellipsoid invariant, and that we may find such a transformation taking any point of this ellipsoid into any other point. Furthermore, if one mass is kept fixed, changes in the ratios of the remaining three will not affect its position. That is

THEOREM IX. *If one of the mass-points of a minimum set has given mass, its locus is an ellipsoid similar to and similarly situated with respect to the equivalent ellipsoid.*

When the equivalent ellipsoid is the unit sphere, from (18) we have

$$\sum_i m_i x_i^2 = M/5 = (sm_1)s(m_2 + m_3 + m_4)M,$$

so that m_1 and the reflection of the foot of the altitude from m_1 are inverse points in a sphere of radius $1/\sqrt{5}$. By applying the transformation U^{-1}, which preserves polar relations, we have:

THEOREM X. *If one of the mass-points of a minimum set is fixed in position, the plane of the other three is the polar with respect to the ellipsoid* (5), $k = 5$, *of the reflection of this point in the origin.*

The reflection may be dispensed with by using the imaginary ellipsoid, $k = -5$, already mentioned in section 2. The fact that the plane is fixed has already been shown in section 3.

By going back to the sphere, or by combination of the last two theorems, we obtain:

THEOREM XI. *If the sum of three of the masses of a minimum set is given, the envelope of the plane of the corresponding mass-points is an ellipsoid similar to, and similarly situated with respect to the equivalent ellipsoid.*

When the equivalent ellipsoid is the unit sphere, we find for the square of the perpendicular from the origin to the line joining m_3 and m_4:

$$s^2m_1^2 + t^2m_2^2 = \frac{m_1 + m_2}{5(m_3 + m_4)},$$

from (18). Thus if $m_3 + m_4$, and hence $m_1 + m_2$, is fixed, the line of m_3 and m_4 is tangent to a sphere, and by a suitable rotation may be turned into any such line. An application of U^{-1} proves:

THEOREM XII. *If the sum of two of the masses of a minimum set is given, the lines of the corresponding pair of mass-points are the tangents to an ellipsoid similar to, and similarly situated with respect to the equivalent ellipsoid.*

The last four theorems deal with minimum sets of four points. For plane distributions, where there are three points in the minimum set, three analogous theorems may be formulated.

5. **Further Conditions for Equimomental Sets.** For completeness, we add a few further conditions for two sets to be equimomental. We begin with two sets which have the same moments of inertia for three distinct co-planor lines through a point. As the plane section of the ellipsoid of inertia for the point in question by the plane of these lines has its center and three radii given, it is determined, and the two systems are equimomental for all axes in this plane. When the systems have the same mass and center of gravity, we may use parallel axes, and so may state:

THEOREM XIII. *Two systems with the same total mass and center of gravity, and the same moments of inertia for the three sides of a triangle, are equimomental for all axes in the plane of this triangle.*

COROLLARY. *If the systems are plane, and the triangle is in their plane, they are equimomental for all axes in space.*

We shall also prove:

THEOREM XIV. *Two systems with the same total mass and*

center of gravity, and the same moments of inertia for the six edges of a tetrahedron, are equimomental.

For, by the preceding theorem, they have the same moments of inertia for all axes in any of the planes of the faces of the tetrahedron, and also for the parallel planes through the center of gravity. Since their ellipsoids of inertia have three distinct central sections identical, they coincide.

These theorems, together with principles of symmetry,[6] such as that a plane or axis of symmetry is necessarily a principal plane or axis of inertia for the center of gravity, frequently simplify the determination of minimum or other sets of points equimomental to a given system.

6. **Illustrations.** The earliest, and perhaps best known[7] example of a minimum set is that for a homogeneous triangular laminor. Such a triangle, of mass M, is equimomental to three particles, each of mass $M/3$ located at the midpoints of the sides. It is also equimomental to four particles, three at the vertices, each of mass $M/12$ and the fourth at the center of gravity, of mass $3M/4$. This is easily verified, either by calculating the moment of inertia for a side, and using the corollary to theorem XIII; or by verifying for an equilateral triangle, which symmetry shows is equimomental to a circle, and applying theorem VI.

For most figures, no one minimum set will be singled out by coördinates simpler than the rest. As a possible set for a sector of a circle, with angle at the vertex A, and radius r, homogeneous of mass M, we mention that composed of one mass

$$M\left(1-\frac{16}{9P}\right) \qquad \text{where} \qquad P=\frac{A}{2\sin\frac{A}{2}}\left(\frac{A}{2\sin\frac{A}{2}}+\cos\frac{A}{2}\right)$$

at the vertex, and two masses, each equal to $8M/9P$, symmetrically situated with respect to the axis of the sector, the lines joining them with its vertex being of length

$$\frac{3r\sqrt{2P}}{8} \text{ and making an angle } \cos^{-1}\left(\frac{\sin A}{A}\right).$$

[6] Appell, l. c.
[7] Routh, l. c.

An equimomental minimum set for a tetrahedron of mass M, is obtained by taking four particles each of mass $M/4$, forming a tetrahedron $1/\sqrt{5}$ times the given one and similar to it and similarly placed, the center of similitude being the common center of gravity of the two. A tetrahedron is also equimomental to five particles, four at the vertices, each of mass $M/20$, and the fifth at the center of gravity, of mass $4M/5$.

As a further illustration we note that a cone of revolution, with altitude h, radius of base a, and mass M, is equimomental to four particles, one of mass $M/16$ at the vertex, and three each of mass $5M/16$ at the vertices of an equilateral triangle whose plane is perpendicular to the axis of the cone, at distance $4h/5$ from the vertex, and the radius of whose circumcircle is $2\sqrt{2}a/5$.

The replacement of systems of masses by equimomental sets of points, minimum or other sets, is particularly useful when the moments of inertia of figures are obtained by dissecting them into simpler figures.[8]

[8] Cf. *American Mathematical Monthly*, vol. 33, 1926, p. 257, 259.

HYPERSURFACES FOR WHICH THE CODAZZI EQUATIONS TAKE THE SIMPLEST FORM

By C. L. E. Moore

1. In a space of constant curvature the equations of Codazzi take the simple form

$$b_{rst} - b_{rts} = 0 \tag{1}$$

where b_{rs} are the coefficients of the second fundamental form and b_{rst} are the derivatives with respect to the element of arc of the hypersurface. Many properties of hypersurfaces in a space of constant curvature depend on this simple form of the equation. It is the purpose of this note to investigate some of the properties of hypersurfaces in a more general space for which the same simple equation holds.

2. We wish to investigate the case when equation (1) is satisfied by a one parameter family of hypersurfaces. Choose the orthogonal trajectories of the family as one set of parameter lines (x^n say). Then the element of arc of V_n is

$$ds^2 = a_{rs}dx^r dx_s, \quad r, s = 1, 2, \cdots, n \tag{2}$$

and that of the hypersurface is

$$ds^2 = a'^{rs}dx^r dx^s, \quad r, s = 1, 2, \cdots, n-1. \tag{3}$$

Let b_{rs} be the components of the second fundamental tensor Choose an orthogonal n-tuple $\lambda_{i|}$ in which to express all tensors and let $\lambda_{n|}$ be the congruence of orthogonal trojectories of the hypersurfaces. Then as usual we can write

$$a_{rs} = \sum_1^n \lambda_{i|r}\lambda_{i|s},$$

$$a_{rs}' = \sum_1^{n-1} \lambda_{i|r}\lambda_{i|s}.$$

The hypersurface component of a tensor $X_{rst}\cdots$ is obtained by contracting it with

$$\Delta^{rst\cdots}_{\alpha\beta\gamma\cdots} = \sum_{i=1}^{n-1} \lambda_{i|\alpha}\lambda_{i|}{}^{r} \sum_{j=1}^{n-1} \lambda_{j|\beta}\lambda_{j|}{}^{s} \sum_{k=1}^{n-1} \lambda_{k|\gamma}\lambda_{k|}{}^{t} \cdots.$$

The Gauss-Codazzi equations can then be written,[1]

$$B_{rstu} = b_{rt}b_{su} - b_{ru}b_{st} = R_{rstu}' - R_{rstu}'' \tag{4}$$

$$b_{rst} - b_{rts} = \Delta^{\alpha\beta\gamma}_{rst} R_{\alpha\beta\gamma\delta}\lambda_{n|}{}^{\delta} \tag{5}$$

where R_{rstu} are the Riemann symbols formed from (2) R_{rstu}' are those formed from (3) and R_{rstu}'' are the hypersurface components of R_{rstu}. It is known that if for every hypersurface (5) takes the form (1), the space V_n must be of constant curvature.

In terms of the orthogonal n-tuple the Riemann tensor takes the form

$$R_{rstu} = \sum_{hijk} \gamma_{hijk}\lambda_{h|r}\lambda_{i|s}\lambda_{j|t}\lambda_{k|u},$$

where γ_{hijk} satisfies the identities [2]

$$\gamma_{hijk} = -\gamma_{hikj} = \gamma_{jkhi}.$$

From these identities it is evident that if (5) takes the form (1) it is necessary and sufficient that

$$\gamma_{hijn} = 0, \quad h \neq n,\ i \neq n. \tag{6}$$

The Ricci tensor is

$$R_{rs} = \sum_{h,\,i,\,j} \gamma_{hihj}\lambda_{i|r}\lambda_{j|s}$$

and the Ricci principal directions are defined by

$$\Big(\sum_{h,\,i,\,j} \gamma_{hihj}\lambda_{i|r}\lambda_{j|s} - \rho a_{rs}\Big)\lambda^{s} = 0.$$

From this it is seen that if

$$\rho = \sum_{h} \gamma_{hnhn},$$

the congruence $\lambda_{n|}$ is one solution. Hence, *if V_n contains a one*

[1] Struik, "Mehrdimensionalen Differentialgeometrie," pp. 124, 135.

[2] Ricci et Levi-Civita, *Methodes de calcul differentiel absolu. Math.*, Ann. Vol. 54, page 157.

parameter family of hypersurfaces for which the Codazzi equations take the form (1), *then the congruence normal to these hypersurfaces is a Ricci principal congruence.*

It is evident that if there is a single hypersurface of this kind then the normals to this surface are Ricci principal directions.

2. The tensor R' satisfies the Bianchi identity

$$R_{rstuv}' + R_{rstu}' + R_{rsuvt}' = 0$$

where the derivatives are taken with respect to the form (3). If (1) is satisfied it is easily seen that B_{rstu} satisfies the same identity,

$$B_{rstuv} + B_{rstu} + B_{rsuvt} = 0,$$

and consequently from (3)

$$R_{rstuv}'' + R_{rstu}'' + R_{rsuvt}'' = 0$$

that is, the hypersurface component of the Riemann tensor of V_n satisfies the Bianchi identity when differentiated with respect to a_{rs}'. If the Riemannian curvature of V_n at each point of the hypersurface and for each orientation lying in it is constant,

$$R_{\alpha\beta\gamma\delta}'' = \sum_{i,j=1}^{n-1} \gamma_{ij} \lambda_{i\alpha} \lambda_{j\beta} \lambda_{i\gamma} \lambda_{j\delta} = K_0(a_{\alpha\gamma}' a_{\beta\delta}' - a_{\alpha\delta}' a_{\beta\gamma}').$$

Now applying the Bianchi identity [3] we obtain the following form of Schur's theorem: *In a space V_n if* (1) *is satisfied for a hypersurface and if at each point of such a hypersurface and for each orientation lying in it the Riemannian curvature of V_n is constant, then the constant is the same for all points of the hypersurface.*

3. The coefficients of the second fundamental form can be written

$$b_{rs} = \sum_{i,j=1}^{n-1} \omega_{ij} \lambda_{i\gamma} \lambda_{j\delta}$$

and making use of this the Codazzi equation can be written in the following form [4]

$$\frac{\partial \omega_{mp}}{\partial s_k} - \frac{\partial \omega_{mk}}{\partial s_p} = \sum_j \omega_{jm}(\gamma_{jkp} - \gamma_{jpk}) + \sum_j (\omega_{jk}\gamma_{jmp} - \omega_{jp}\gamma_{jmk}).$$

[3] Bianchi, "Geometria differenziale," 2d Ed.; Vol. 1, page 351.
[4] Ricci et Levi-Civita, loc. cit., page 170.

For a congruence of lines of curvature $\omega_{ij} = 0$, $i \neq j$. Then for such a congruence, putting $p = m$, we have

$$\frac{\partial \omega_{mm}}{\partial s_k} = (\omega_{kk} - \omega_{mm})\gamma_{kmm}.$$

Then for hypersurfaces for which (1) is satisfied we have

(a) *If a congruence of lines of curvature is also geodesic* ($\gamma_{nmm} = 0$) *then the curvature of the lines of this congruence at* δ *points on another line of curvature is constant.*

(b) *If* $\omega_{mm} = 0$, $\omega_{kk} \neq 0$, *then* $\gamma_{kmm} = 0$ *and the lines of curvature are geodesics.*

(c) *If the curvatures* ω_{ii} *are all equal to all points, they are constant*

If we choose the directions of the lines of curvature as the orthogonal n-tuple

$$b_{rs} = \sum_i \omega_{ii}\lambda_{i|r}\lambda_{i|s}$$

and consequently

$$b = |b_{rs}| = |\sum \omega_{ii}\lambda_{i|r}\lambda_{i|s}| = \omega_{11}\omega_{22}\cdots\omega_{nn}|a_{rs}|$$

or

$$\frac{b}{a} = \omega_{11}\omega_{22}\cdots\omega_{nn} = \omega.$$

If ξ^a is the unit normal vector to the hypersurface[5]

$$\frac{d\xi^a}{ds_i} = -\omega_{ii}\lambda_{i|}{}^a.$$

Now if for the congruence λ_i, $\omega_{ii} = 0$, we see that the normals to the surface along a line of this congruence are all parallel, these lines are then asymptotes. If an asymptotic line of a hypersurface is also a geodesic of the hypersurface, it is also a geodesic of the enveloping space V_n. We see from (b) that these lines are geodesics of V_n. Hence, *if equation* (1) *is satisfied and* $\omega = 0$, *the hypersurface has a geodesic of* V_n *passing through each point and the tangent planes at points of such a geodesic are all parallel. Conversely, if a hypersurface contains an* $(n-2)$ *parametered family of lines, such that the tangent hyperplanes*

[5] Eisenhart, "Riemannian Geometry," page 173, formula (51, 19).

along any one of them are parallel, then if (1) *is satisfied the lines will be geodesics of* V_n.

The generalization for the vanishing of any number of the coefficients ω_{ii} is evident.

In case the V_n is a V_3 the above theorem becomes. *If equation* (1) *holds and the determinant of the second fundamental form of a surface vanishes, then the surface contains a pencil of geodesics.*

If $n = 3$ the Riemann symbols can be written in the form [6]

$$\alpha_{rs} = \gamma_{hk}\lambda_{h|r}\lambda_{k|s}$$

and the principal directions of α_{rs} which coincide with the Ricci principal directions [7] are given by

$$(\gamma_{hk}\lambda_{h|r}\lambda_{k|s} - \rho a_{rs})\lambda^r = 0.$$

It is at once evident that the necessary and sufficient condition that $\lambda_{3|}$ be a principal congruence is

$$\gamma_{3h} = 0, \quad h \neq 3.$$

If we take $\lambda_{1|}$ and $\lambda_{2|}$ as principal congruences, the Riemann symbol takes the form

$$\alpha_{rs} = \sum \gamma_{hk}\lambda_{h|r}\lambda_{k|s}.$$

Now if $\lambda_{3|}$ is a normal congruence, the surfaces perpendicular to it will have Codazzi equations of the form (1). Hence; *In a Riemannian space* V_3, *the necessary and sufficient condition that there exist a set of surfaces whose Codazzi equations are of form* (1) *is that one principal congruence be normal.*

For such a V_3 (1) can be written [8]

$$\frac{\partial b_{11}}{\partial x^2} - \frac{\partial b_{12}}{\partial x^1} - \left\{\begin{matrix}1 & 2\\ 1\end{matrix}\right\} b_{11} + \left[\left\{\begin{matrix}1 & 1\\ 1\end{matrix}\right\} - \left\{\begin{matrix}1 & 2\\ 2\end{matrix}\right\}\right] b_{12} + \left\{\begin{matrix}1 & 1\\ 2\end{matrix}\right\} b_{22} = 0,$$

$$\frac{\partial b_{22}}{\partial x^1} - \frac{\partial b_{12}}{\partial x^2} - \left\{\begin{matrix}2 & 2\\ 1\end{matrix}\right\} b_{11} + \left[\left\{\begin{matrix}2 & 2\\ 2\end{matrix}\right\} - \left\{\begin{matrix}1 & 2\\ 1\end{matrix}\right\}\right] b_{12} + \left\{\begin{matrix}1 & 2\\ 2\end{matrix}\right\} b_{22} = 0.$$

[6] Ricci et Levi-Civita, loc. cit.
[7] Ricci, Direzioni e invarianti principali, *Atti Veneto*, Vol. 63, pp. 1233–1239.
[8] Bianchi, "Geometria Differenziale," 2d Ed., page 491.

or

$$(7)\quad \begin{cases} \dfrac{\partial}{\partial x^2}\left(\dfrac{b_{11}}{\sqrt{a}}\right) - \dfrac{\alpha}{\alpha x^1}\left(\dfrac{b_{12}}{\sqrt{a}}\right) + \left\{\begin{matrix}2 & 2\\ 2\end{matrix}\right\}\dfrac{b_{11}}{\sqrt{a}} - 2\left\{\begin{matrix}1 & 2\\ 2\end{matrix}\right\}\dfrac{b_{12}}{\sqrt{a}} + \left\{\begin{matrix}1 & 1\\ 2\end{matrix}\right\}\dfrac{b_{22}}{\sqrt{a}} = 0, \\ \dfrac{\partial}{\partial x^1}\left(\dfrac{b_{22}}{\sqrt{a}}\right) - \dfrac{\partial}{\partial x^2}\left(\dfrac{b_{12}}{\sqrt{a}}\right) + \left\{\begin{matrix}2 & 2\\ 1\end{matrix}\right\}\dfrac{b_{11}}{\sqrt{a}} - 2\left\{\begin{matrix}1 & 2\\ 1\end{matrix}\right\}\dfrac{b_{12}}{\sqrt{a}} + \left\{\begin{matrix}1 & 1\\ 1\end{matrix}\right\}\dfrac{b_{22}}{\sqrt{a}} = 0. \end{cases}$$

The Gauss equation takes the form

$$k = \frac{b}{a} = K - K_0$$

where K is the Gaussian curvature of the surface and K_0 Riemannian curvature with respect to V_3 and k is the relative curvature.

If now we consider a surface of negative relative curvature

$$k = \frac{1}{\rho^2}$$

$$b_{11} = b_{22} = 0, \qquad \frac{b_{12}}{\sqrt{a}} = \frac{1}{\rho}$$

and equations (7) become

$$\frac{\partial \log \rho}{\partial x^1} = 2\left\{\begin{matrix}1 & 2\\ 2\end{matrix}\right\}, \qquad \frac{\partial \log \rho}{\partial x^2} = 2\left\{\begin{matrix}1 & 2\\ 1\end{matrix}\right\}.$$

For a surface of constant relative curvature

$$\left\{\begin{matrix}1 & 2\\ 2\end{matrix}\right\} = \left\{\begin{matrix}1 & 2\\ 1\end{matrix}\right\} = 0$$

and consequently a_{11} is a function of x_1 only and a_{22} a function of x_2 only. Making change of variables the element of arc can then be put in the form

$$(8)\qquad ds^2 = (dx^1)^2 + 2\cos\omega dx^1 dx^2 + (dx^2)^2.$$

If the surface has zero curvature, that is $K = 0$, and also k constant this would require that $K_0 = \text{const.}$ In this case the condition that the surface whose element of arc is (8) have

curvature zero is

$$\frac{\partial^2 \omega}{\partial x^1 \partial x^2} = 0,$$

and consequently (8) can be written

$$ds^2 = (dx_1)^2 + (U + V)dx^1 dx^2 + (dx^2)^2$$

where U is a function of x^1 only and V a function x^2 only. Now if the tangent to the asymptote x^1 is moved along the asymptote x^2. That is, moving the vector $(dx^1, 0)$ in the direction $(0, dx_2)$

$$\delta_\mu dx^1 = \sum \left\{ \begin{matrix} i & j \\ k & \end{matrix} \right\} dx^i dx^j = 0$$

and consequently the asymptotes are parallel according to Levi-Civita.

If one of the set of asymptotic lines are also geodesics, then they must be parallel.

HARMONIC ANALYSIS AND GROUP THEORY

By Norbert Wiener

Weyl [1] has developed in some detail the relations between the theory of unitary groups and the theory of periodic and almost periodic functions. The groups which he introduces are one-parameter groups of linear functional transformations which leave invariant the expression

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}f(x)\bar{f}(x)dx \tag{1}$$

(which we shall abbreviate $M_x(f(x)\bar{f}(x))$. A similar theory of functions expansible in Fourier integral form would involve the invariancy of the integral quadratic form

$$\int_{-\infty}^{\infty}f(x)\bar{f}(x)dx. \tag{2}$$

The author [2] has shown that the theory of almost periodic functions may very satisfactorily be established on the basis of the Fourier integral. If this treatment is to be associated with the theory of groups, it becomes necessary to consider transformations which leave both (1) and (2) invariant.

If (1) exists, the author has shown [3] that

$$\left[\int_{A}^{\infty}+\int_{-\infty}^{-A}\right]\left|\frac{f(x)}{x}\right|^2dx$$

is finite. In what follows, we shall consider only measurable functions fulfilling this condition, and such that

$$\int_{-A}^{A}|f(x)|^2dx$$

is finite. All integrals are of course understood in the sense of Lebesgue.

[1] H. Weyl, "Integralgleichungen und fastperiodische Funktionen," *Math. Annalen.*, 97 (1926), 338–356.

[2] N. Wiener, "The Spectrum of an Arbitrary Function," *Proc. Lond. Math. Soc.*, Ser. 2, 27, 483–496.

[3] N. Wiener, "On the representation of Functions by Trigonometrical Integrals," *Math. Ztschr.* 24 (1925), 577.

Let T be a transformation defined for all functions of this class, and preserving (2) invariant for all the functions for which it is defined. Let S_A be the transformation leading from $f(x)$ to the function $f_A(x)$, where

$$f_A(x) = \begin{cases} 0 \text{ if } |x| > A \\ f(x) \text{ if } |x| \leq A \end{cases}.$$

There will be certain transformations T keeping (2) invariant and such that for all functions f of our class for which $M_x(|f(x)|^2)$ exists,

$$\int_{-\infty}^{\infty} |(TS_A - S_AT)f(x)|^2 dx = o(A). \tag{3}$$

For such a transformation, we may readily deduce from

$$\lim_{A\to\infty} \frac{1}{2}\int_{-\infty}^{\infty} \left|\frac{TS_Af(x)}{\sqrt{A}}\right|^2 dx = M_x(|f(x)|^2)$$

and

$$\lim_{A\to\infty} \int_{-\infty}^{\infty} \left|\frac{TS_Af(x)}{\sqrt{A}} - \frac{S_ATf(x)}{\sqrt{A}}\right|^2 dx = 0$$

that

$$M_x(|Tf(x)|^2) = \lim_{A\to\infty} \frac{1}{2}\int_{-\infty}^{\infty} \left|\frac{S_ATf(x)}{\sqrt{A}}\right|^2 dx = M_x(|f(x)|^2).$$

These transformations form a group, for

$$\begin{aligned}
&\int_{-\infty}^{\infty} |(T_1T_2S_A - S_AT_1T_2)f(x)|^2 dx \\
&= \int_{-\infty}^{\infty} |[(T_1T_2S_A - T_1S_AT_2) + (T_1S_AT_2 - S_AT_1T_2)]f(x)|^2 dx \\
&\leq 2\int_{-\infty}^{\infty} |(T_1T_2S_A - T_1S_AT_2)f(x)|^2 dx \\
&\quad + 2\int_{-\infty}^{\infty} |(T_1S_AT_2 - S_AT_1T_2)f(x)|^2 dx \\
&= 2\int_{-\infty}^{\infty} |(T_2S_A - S_AT_2)f(x)|^2 dx \\
&\quad + 2\int_{-\infty}^{\infty} |(T_1S_A - S_AT_1)T_2f(x)|^2 dx.
\end{aligned}$$

Furthermore,

$$\int_{-\infty}^{\infty} |(T^{-1}S_A - S_AT^{-1})f(x)|^2dx$$
$$= \int_{-\infty}^{\infty} |(S_A - TS_AT^{-1})f(x)|^2dx$$
$$= \int_{-\infty}^{\infty} |(TS_A - S_AT)T^{-1}f(x)|^2dx.$$

Thus the product of two transformations satisfying (3) and the inverse of a transformation satisfying (3) also satisfy (3).

An example of a transformation satisfying (3) is

$$T^\lambda f(x) = f(x + \lambda),$$

for

$$\int_{-\infty}^{\infty} |(T^\lambda S_A - S_AT^\lambda)f(x)|^2dx$$
$$= \int_0^\lambda [|f(x - A - \lambda)^2 + f(x + A - \lambda)|^2]dx = o(A).$$

We shall call a transformation satisfying (3) and keeping (2) invariant *properly unitary*. Let us consider the one-parameter group of properly unitary transformations U^λ, where

$$U^{\lambda+\mu} = U^\lambda U^\mu.$$

Let $f(x)$ be such that

$$g(t) = M_x[(U^tf(x))\bar{f}(x)]. \tag{4}$$

exists for every t. Clearly by the Schwarz inequality

$$g(t) \leq \sqrt{M_x(|U^tf(x)|^2)M_x(|f(x)|^2)}$$
$$= \sqrt{g(0)\lim_{A\to\infty}\frac{1}{2A}\int_{-\infty}^{\infty}|S_AU^tf(x)|^2dx}$$
$$= \sqrt{g(0)\lim_{A\to\infty}\frac{1}{2A}\int_{-\infty}^{\infty}[U^tS_Af(x)|^2dx}$$
$$= g(0).$$

It follows that $g(t)$ is a bounded function, and that

$$R(u_1, u_2) = \text{l.i.m.}_{B\to\infty} \frac{1}{\pi} \int_{-B}^{B} g(t) \frac{\sin (u_2 - u_1)t/2}{t} e^{i(u_1+u_2/2)t} dt$$

exists when $(u_1 - u_2)/2$ is given as a quadratically summable function of $(u_1 + u_2)/2$. For this one only need apply Plancherel's theorem.[4]

Let us put

$$g_A(t) = \frac{1}{2A} \int_{-\infty}^{\infty} [U^t S_A f(x)] \overline{S_A f(x)} dx. \tag{5}$$

Under condition (4) for every t, we have

$$g(t) = \lim_{A\to\infty} \frac{1}{2A} \int_{-\infty}^{\infty} [S_A U^t f(x)] \overline{S_A f(x)} dx,$$

which one readily reduces to the form

$$g(t) = \lim_{A\to\infty} g_A(t)$$

by means of (3). We may easily prove $g_A(t)$ to be less than or equal to $g_A(0)$, and hence to be uniformly bounded in t and A, as $g_A(0) \to g(0)$ as $A \to \infty$.

Let us now introduce a new assumption concerning U^t. Let the transformation W taking $f(x)$ into $\varphi(a, t)$ (a fixed), where

$$\varphi(x, t) = U^t f(x)$$

preserve (2) invariant. Then if $\psi(x) = WS_A f(x)$,

$$g_A(t) = \frac{1}{2A} \int_{-\infty}^{\infty} \psi(x + t) \bar{\psi}(x) dx.$$

It follows readily that $g_A(t)$ is absolutely integrable over $(-\infty, \infty)$ and[5] that

$$P_A(u) = \frac{1}{\pi} \int_{-\infty}^{\infty} g_A(t) e^{iut} dt$$

[4] Hobson, "The Theory of Functions of a Real Variable," 2d edition 2 (1926), 742–752.

[5] N. Wiener, "The Spectrum of an Arbitrary Function," p. 483.

exists and is real and positive.[6] Thus

$$R_A(u_1, u_2) = \frac{1}{\pi}\int_{-\infty}^{\infty} g_A(t)\,\frac{\sin\,(u_2 - u_1)t/2}{t}\,e^{i(u_1+u_2/2)t}dt$$

exists, is monotone in u_1 and u_2, and has the property that

$$R_A(u_1, u_2) + R_A(u_2, u_3) = R_A(u_1, u_3).$$

To see this, we need only reflect that

$$R_A(u_1, u_2) = \int_{u_1}^{u_2} P_A(u)du.$$

Now

$$g(t)\,\frac{\sin\,(u_2 - u_1)t/2}{t} = \underset{A\to\infty}{\text{l.i.m.}}\, g_A(t)\,\frac{\sin\,(u_2 - u_1)t/2}{t},$$

so that

$$R(u_1, u_2) = \underset{A\to\infty}{\text{l.i.m.}}\, R_A(u_1, u_2).$$

From this we may readily conclude[7] that

$$R(u_1, u_2) + R(u_2, u_3) = R(u_1, u_3),$$

and that $R(u_1, u_2)$ may be so defined as to be monotone in both arguments and increasing in u_2.

We have

$$R_A(-u, u) = \frac{1}{\pi}\int_{-\infty}^{\infty} g_A(t)\,\frac{\sin ut}{t}\,dt.$$

Hence

$$\frac{1}{u}\int_0^u R_A(-u, u)du = \frac{1}{\pi}\int_{-\infty}^{\infty} g_A(t)\,\frac{1-\cos ut}{t^2}\,dt.$$

Now,

$$R_A(-\infty, \overset{\prime}{\infty}) = \lim_{u\to\infty}\frac{1}{u}\int_0^u R_A(-u, u)du$$

$$= \lim_{A\to\infty}\frac{1}{\pi u}\int_{-\infty}^{\infty} g_A(t)\,\frac{1-\cos ut}{t^2}\,dt = g_A(0).$$

[6] Indeed, we might assume directly that $\frac{1}{\pi}\int_{-\infty}^{\infty} e^{iut}dt\int_{-\infty}^{\infty}[U^t f(x)]\bar{f}(x)dx$ is positive definite and exists for every quadratically summable $f(x)$.

[7] N. Wiener, loc. cit., p. 489.

by a Fejer theorem. Hence in the limit,

$$R(-\infty, \infty) \leq g(0).$$

In the same way, we can show[8] that

$$\frac{1}{u}\int_0^u R(-u, u)du = \frac{1}{\pi u}\int_{-\infty}^{\infty} g(t)\frac{1-\cos ut}{t^2}dt.$$

Hence

$$R(-\infty, \infty) = \lim_{u\to\infty}\frac{1}{\pi u}\int_{-\infty}^{\infty} g(t)\frac{1-\cos ut}{t^2}dt.$$

It follows by a Tauberian theorem[9] that

$$R(-\infty, \infty) = \lim_{\epsilon\to 0}\frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} g(t)dt.$$

The theory of almost periodic functions may also be generalized to more general groups of transformations than the translation group. It is an interesting fact that the conditions on this group are somewhat different from those of spectrum theory. This is connected with the fact that, as Franklin has shown,[10] all of the theory of almost periodic functions and of their extensions follows by elementary means from the classical Bohr approximation theorem for almost periodic functions of the usual sort.

Let $\{T^\lambda\}$ be a set of linear transformations into itself of the class of all functions $f(x)$ of summable square modulus over any finite interval, and such that $\frac{1}{A}\int_{-A}^{A}|f(x)|^2dx$ is bounded in A. Let

$$T^\lambda(T^\mu(f(x))) = T^{\lambda+\mu}(f(x)).$$

Let the transformations T^λ be unitary, and let

$$M_\lambda T^\mu(f(x, \lambda)) = T^\mu M_\lambda f(x, \lambda).$$

We shall say that the function $f(x)$ has a T transformation-number τ_ϵ with respect to ϵ when for all x and λ,

$$|T^{\tau_\epsilon}f(x) - f(x)| < \epsilon.$$

[8] Ibid., p. 489.

[9] Cf. N. Wiener, "A New Method in Tauberian theorems," this *Journal*, vol. 6 (October, 1928), p. 163.

[10] P. Franklin, *Math. Zeitschr.*, 29 (1928), pp. 70–86.

In case for every $\epsilon > 0$ there is an L_ϵ such that for every mod A, $(A, A + L_\epsilon)$ contains at least one T transformation-number τ_ϵ, and in case $T^\lambda f(x)$ is equally continuous in λ for all x we shall say that $f(x)$ is almost periodic with respect to the group $\{T^\lambda\}$.

The arguments of the author's previous paper [11] hold substantially without change to show that we can approximate uniformly in λ and x to $T^\lambda f(x)$ by polynomials of the form $\sum_1^N A_n(x)e^{i\lambda_n\lambda}$. These polynomials may be obtained from $f(x)$ by a series of steps from f_{k-1} to f_k of the form

$$\begin{aligned} f_k(x, \lambda) &= M_\lambda\{T^\lambda f_{k-1}(x, 0)K(\lambda - \mu)\} \\ &= M_\lambda\{T^{\mu+\lambda} f_{k-1}(x, 0)K(\lambda)\} \\ &= T^\mu M_\lambda\{T^\lambda f_{k-1}(x, 0)K(\lambda)\}, \end{aligned}$$

and by a similar process, we can write

$$\begin{aligned} A_n(x)e^{i\lambda_n\mu} &= M_\lambda\{T^\lambda[\sum_1^N A_n(x)]e^{-i\lambda_n(\lambda-\mu)}\} \\ &= T^\mu M_\lambda\{T^\lambda[\sum_1^N A_n(x)]e^{-i\lambda_n\lambda}\}. \end{aligned}$$

Letting $\mu = 0$,

$$A_n(x) = M_\lambda\{T^\lambda[\sum_1^N A_n(x)]e^{-i\lambda_n\lambda}\},$$

and

$$A_n(x)e^{i\lambda_n\mu} = T^\mu A_n(x),$$

or in particular,

$$TA_n(x) = e^{i\lambda_n}A_n(x).$$

Thus $e^{i\lambda_n}$ is an Eigenwert of T possessing $A_n(x)$ as an Eigenfunktion, and $f(x)$ may be approximated uniformly by the sum $\sum_1^N A_n(x)$ of Eigenfunktionen of the transformation T.

[11] "The Spectrum of an Arbitrary Function," 494–6.

NOTE ON THE STATISTICAL INTERPRETATION OF MAXWELL'S EQUATIONS *

By M. S. Vallarta

It has often been trivially noted in the literature that Schrödinger's relativistic equation in a force-free field [1] reduces for $m = 0$ to the equation of propagation of light in free space. In accordance with the general principles of quantum mechanics Schrödinger's equation is obtained from the energy-momentum relation $-p_0^2 + p_x^2 + p_y^2 + p_z^2 + m^2c^2 = 0$ by substituting for the components of the four-vector p the operators $p_0 = (h/2\pi ic)\partial/\partial t$, $p_x = (h/2\pi i)\partial/\partial x$, etc. For a photon the proper mass vanishes and the coefficients $h/2\pi i$ divide out of the corresponding wave equation. In this sense the equation of propagation of light can be regarded as the wave equation of the photon just as Schrödinger's equation is the wave equation of the electron without spin.[2]

In the statistical interpretation of the non-relativistic wave function, which alone is consistent with the duality between particle picture and wave picture so ably discussed by Heisenberg and Bohr, the quantity $\bar{\psi}\psi$ gives the probability of an electron being at an assigned point (x, y, z). This interpretation, thoroughly satisfactory from the point of view of observed physical facts, depends essentially upon the existence of normal and orthogonal characteristic functions satisfying the non-relativity equation, which existence in turn depends upon the fact that the latter is linear in the energy E. This is not true of the Schrödinger relativistic equation when there is an

* Presented to the *American Physical Society*, Washington, April 18, 1929.

[1] E. Schrödinger, Ann. d. Phys. 81, 139, 1926. See Equation (36), p. 163, with $A = 0$ and $V = 0$.

[2] Compare H. Weyl, Gruppentheorie und Quantenmechanik, p. 46, Leipzig, 1928.

electromagnetic field present, and for a while, in spite of the efforts of Gordon and Klein, it remained a serious difficulty in quantum mechanics. It was first removed by the brilliant work of Dirac who as a matter of fact starts from the requirement that the relativistic wave equation of the electron shall be linear in the energy.[3]

If one considers the duality between photon and electromagnetic wave in the same way as the duality between electron and De Broglie wave, the question arises as to whether any such probability considerations as apply to the Schrödinger wave function are also applicable to the electromagnetic vectors. In other words, is it possible to interpret the Maxwell equations in a statistical sense? The results of the present note show that such an interpretation is possible at least for those electromagnetic quantities which like the energy density and the Poynting vector depend quadratically upon the electric and the magnetic vectors. Since, except at low frequencies, none except these quantities are experimentally accessible, they alone can have a direct physical meaning.[4]

If we introduce the complex vector $M = E + iH$ ($i = \sqrt{-1}$), where E and H are the electric and magnetic field vectors, Maxwell's equations in free space are written in ordinary Gibbs' vector notation

$$i\nabla \times M = -\frac{1}{c}\frac{\partial M}{\partial t}, \qquad i\nabla \cdot M = 0.$$

Assuming now a dummy function $M_4 = 0$ (or const.) these equations can be written

$$(p_0 + \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3)M = 0, \tag{1}$$

where the p's have the same meaning as before (note that the common factor $h/2\pi i$ divides out) and the β's are four-rowed

[3] P. A. M. Dirac, Proc. Royal Soc. of London, 117, p. 610, 1928; 118, p. 351, 1928, hereinafter referred to as "Dirac I" and "II." See also Dirac's Leipzig lecture in "Quantentheorie und Chemie," p. 85 (collected by Falkenhagen, Leipzig, 1928).

[4] Compare Norbert Wiener, Jour. of Franklin Institute, p. 526, April 1929.

Hermitian matrices

$$\beta_1 = \begin{Vmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{Vmatrix}, \qquad \beta_2 = \begin{Vmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \\ -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{Vmatrix},$$

$$\beta_3 = \begin{Vmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{Vmatrix},$$

which satisfy the relations

$$\beta_j\beta_k + \beta_k\beta_j = 0, \qquad \beta_j^2 = 1. \qquad (j \neq k,\ j, k = 1, 2, 3)$$

By multiplying (1) by $-p_0 + \beta_1 p_1 + \beta_2 p_2 + \beta_3 p_3$, one obtains the equation of propagation of light

$$(-p_0^2 + p_1^2 + p_2^2 + p_3^2)M = 0.$$

Equation (1) bears a striking formal resemblance to Dirac's equation of the spinning electron

$$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 mc)\psi = 0$$

when the electron mass m vanishes. Here the p's are the same as above and, with Dirac's choice of the α's:

$$\alpha_1 = \begin{Vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{Vmatrix}, \qquad \alpha_2 = \begin{Vmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{Vmatrix},$$

$$\alpha_3 = \begin{Vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{Vmatrix}, \qquad \alpha_4 = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix},$$

which satisfy the relations

$$\alpha_j\alpha_k + \alpha_k\alpha_j = 0, \qquad \alpha_j^2 = 1. \qquad (j \neq k,\ j, k = 1, 2, 3, 4)$$

Dirac (l. c. I, pp. 616/617) has further shown that all sets of four-rowed matrices satisfying these relations can be obtained from one another by applying a suitable canonical transformation.

In Dirac's theory $\sum \bar{\psi}_i \psi_i$ and $-\sum \bar{\psi}_i (\alpha_i \psi)_i$ are interpreted as (Dirac II, pp. 352/354) the probability per unit volume of the electron being within that volume, and the probability of the electron crossing unit area perpendicular to the three coördinate axes respectively. Dirac further proves that these four quantities transform like a four-vector under a Lorentz transformation, of which vector $\sum \bar{\psi}_i \psi_i$ is the time component, and that they satisfy a continuity equation

$$\frac{1}{c} \frac{\partial}{\partial t} \sum \bar{\psi}_i \psi_i - \sum \frac{\partial}{\partial x_k} \bar{\psi}_i (\alpha_k \psi)_i = 0.$$

They can therefore be interpreted as the electric density and the electric current density, respectively.

In the case of the Maxwell equations written in the form (1), the corresponding quantities are $\bar{M} \cdot M$ and $-\sum \bar{M}_i (\beta_i M)_i$. The diagonal matrix $\beta_0 = 1$ which may be taken as multiplying the operator p_0, and $\beta_1, \beta_2, \beta_3$ are the components of a four-vector, of which β_0 is the time component. $\bar{M} \cdot M = \sum \bar{M}_i (\beta_0 M)_i$ is therefore the time component of a four vector, of which $-\sum \bar{M}_i (\beta_i M)_i$ are the three space components. Our operator in Equation (1) is the Hermitian operator

$$\begin{Vmatrix} \partial/\partial t & -i\partial/\partial z & i\partial/\partial y & -i\partial/\partial x \\ i\partial/\partial z & \partial/\partial t & -i\partial/\partial x & -i\partial/\partial y \\ -i\partial/\partial y & i\partial/\partial x & \partial/\partial t & -i\partial/\partial z \\ i\partial/\partial x & i\partial/\partial y & i\partial/\partial z & \partial/\partial t \end{Vmatrix},$$

but we are not at liberty to assume it to be always Hermitian, for if we apply a Lorentz transformation and divide out by the new coefficient of p_0, the resultant operator will not be Hermitian, although it can be brought back to Hermitianness by a suitable canonical transformation. If, without the assumption of Hermitianness, we consider the transposed operator

$$(-\tilde{\beta}_0 p_0 - \tilde{\beta}_1 p_1 - \tilde{\beta}_2 p_2 - \tilde{\beta}_3 p_3) N = 0, \tag{2}$$

the above quantities are replaced by $\sum N_i (\beta_0 M)_i$ and $-\sum N_i \times (\beta_i M)_i$. Multiply now (1) by N and (2) by M and subtract. We obtain

$$N(\beta_0 p_0 + \sum \beta_i p_i) M + M(\tilde{\beta}_0 p_0 + \sum \tilde{\beta}_i p_i) N = 0,$$

or

$$N\left(\beta_0\frac{\partial}{c\partial t}-\sum\beta_i\frac{\partial}{\partial x_i}\right)M+\frac{1}{c}\frac{\partial N}{\partial t}\beta_0 M-\sum\frac{\partial N}{\partial x_j}\beta_i M=0,$$

and we have immediately the conservation theorem

$$\frac{1}{c}\frac{\partial}{\partial t}\sum N_j(\beta_0 M)_j-\sum\frac{\partial}{\partial x_i}N_j(\beta_i M)_j=0, \tag{3}$$

which for an Hermitian operator reduces to

$$\frac{1}{c}\frac{\partial}{\partial t}\sum\bar{M}_j(\beta_0 M)_j-\sum\frac{\partial}{\partial x_i}\bar{M}_j(\beta_i M)_j=0. \tag{3'}$$

These quantities may therefore be interpreted, as in the Dirac theory, as the probability per unit volume of n photons being within that volume,[5] and the probability of n photons crossing unit area perpendicular to the three axes respectively. The continuity equation expresses then that the change in probability of n photons being within a boundary is equal to the probability of their having crossed the boundary.

In Dirac's theory it is shown (l. c. III) that the matrix

$$\rho_{mn}=\int\sum_i\bar{\psi}_{im}\psi_{in}dv$$

can have only the characteristic values -1 or 0. That is the charge within a closed region can be either a multiple of $-e$ or 0, a theorem which expresses in the most direct way the fundamental fact of the atomicity of electricity. We should expect an analogous theorem in the statistical interpretation here given, which will be discussed below.

The meaning of the components of the four vector found above in the Maxwell theory is immediately found. We have, with $M_4=0$:

$$\begin{aligned}
&\sum\bar{M}_j(\beta_0\mathrm{M})_j=\bar{M}\cdot M=(E-iH)\cdot(E+iH)=E^2+H^2=W,\\
&-\sum\bar{M}_j(\beta_1 M)_j=i(\bar{M}_2M_3-\bar{M}_3M_2)\\
&\qquad\qquad=i(\bar{M}\times M)_x=2(E\times H)_x=KS_x,\\
&-\sum\bar{M}_j(\beta_2 M)_j=i(\bar{M}_3M_1-\bar{M}_1M_3)\\
&\qquad\qquad=i(\bar{M}\times M)_y=2(E\times H)_y=KS_y,\\
&-\sum\bar{M}_j(\beta_3 M)_j=i(\bar{M}_1M_2-\bar{M}_2M_1)\\
&\qquad\qquad=i(\bar{M}\times M)_z=2(E\times H)_z=KS_z,
\end{aligned}$$

[5] Compare J. C. Slater, "Light Quanta and Wave Mechanics," Phys. Rev. 31, p. 895, 1928.

where W is the energy density and S_x, S_y, S_z the components of the Poynting vector. We have thus found the statistical interpretation of these quantities. The continuity equation (3′) is the well-known equation of electromagnetic theory $\nabla \cdot S + \partial W/\partial t = 0$. A difficulty might be encountered in the fact that under a Lorentz transformation the electric density and the electric current transform like a four vector, whereas the energy density and Poynting vector transform like the space-time components of the symmetrical energy-momentum tensor.[6] But it must be remembered that we are dealing here not with the energy and the Poynting vector *at a point*, but per unit volume and unit area. The energy and electromagnetic momentum per unit volume transform like a four-vector (Pauli, l. c. p. 641).[7]

In order to complete the proof of the statistical interpretation of the energy density and the Poynting vector it would be required to show that the energy contained within a closed volume is either $nh\nu$ or zero, where n is an integer and ν may or may not be continuous, corresponding to the fact that the characteristic values of the wave equation form a continuous enumerable set. The proof is difficult on account of the appearance of continuous matrices. For the case of a monochromatic radiation the theorem becomes trivial.

If instead of the Hermitian matrices β_j in Equation (1) we introduce the real matrices $\beta_j' = (1/i)\beta_j$ which satisfy the relations

$$\beta_j'^2 = -1, \qquad \beta_j'\beta_k' + \beta_k'\beta_j' = 0. \qquad (j \neq k, j, k = 1, 2, 3)$$

Maxwell's equations may be written

$$(ip_0 + \beta_1'p_1 + \beta_2'p_2 + \beta_3'p_3)M = 0, \tag{4}$$

and it is immediately seen that the algebra of the operator $p_0 + \beta_1'p_1 + \beta_2'p_2 + \beta_3'p_3$ is precisely quaternion algebra. In-

[6] W. Pauli, "Relativitätstheorie," Enz. d. math. Wiss., Vol. V, 19, p. 639. Compare D. J. Struik and M. S. Vallarta, Jour. of Franklin Institute, p. 501, April, 1929.

[7] I owe this remark to a conversation with Professor Heisenberg.

stead of the operator in Equation (4) we may use the operator [8]

$$\Diamond = \frac{\sqrt{-1}}{c}\frac{\partial}{\partial t} + i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z},$$

where $i, j, k, \sqrt{-1}$ satisfy the multiplication rules of quaternion algebra

$$i^2 = j^2 = k^2 = -1, \qquad (\sqrt{-1})^2 = -1,$$
$$i\sqrt{-1} = \sqrt{-1}i, \qquad j\sqrt{-1} = \sqrt{-1}j, \qquad k\sqrt{-1} = \sqrt{-1}k,$$
$$ij = -ji = k, \qquad jk = -kj = i, \qquad ki = -ik = j.$$

Maxwell's equations are then simply written

$$\Diamond M = 0,$$

and we note that again

$$\bar{\Diamond}\Diamond M = -\nabla^2 M + (1/c^2)\partial^2 M/\partial t^2 = 0.$$

It is a pleasure to acknowledge my indebtedness to the members of the Seminar on Quantum Mechanics for helpful discussions and suggestions.

[8] Compare F. Klein, "Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert," Vol. II, p. 85, Berlin, 1927.

ALMOST PERIODIC FUNCTIONS AND VIBRATING SYSTEMS

By Carl F. Muckenhoupt

Contents

Introduction 163
Chapter I. Theory of functions almost periodic in the mean 166
Chapter II. Proof of the almost periodicity in the mean of a loaded vibrating string 182
Chapter III. Existence of the solution of the problem 190
Chapter IV. The completion of the problem, and the closure of the eigenfunktionen or characteristic functions 195
Bibliographical Appendix 198

Introduction

The theory of almost periodic functions was founded by Harald Bohr [1] in a paper which treated of the functions of one variable. The general theory was then treated by him in two later papers,[2] finally resulting in the theorems for n variables. Other early work was done by Bochner [3] who developed the theory for an infinite number of variables, P. Franklin,[4] H. Weyl,[5] W. Stepanoff [6] who extended the notions to include functions not summable in the ordinary sense, and N. Wiener [7] who developed theorems for almost periodic functions as part of a paper treating of function spectra. This list is not at all intended to be exhaustive, and a more complete bibliography will be found as an appendix, but it serves as an introduction to a type of function which, while the theory has been very fully considered and worked out, has not been definitely applied to physical problems.

Consequently, as a step in this direction it is the intention of the present paper to treat finite vibrating systems of an infinite number of degrees of freedom through the medium of almost periodicity, and so attempt to bring out the advantages and

difficulties of the method, to systematize and generalize the attack, and to classify fields where such proceedure is fruitful.

Specifically, the first problem taken up is that of the vibrating loaded string. An insight into the fundamental workings of such a system is gained from the application of this mode of approach, and the relations of various factors seen, as it were, from a different vantage point. It is hoped that the complexities of procedure will not cloud the insight afforded, when other systems are encountered. However, as will be pointed out later, the difficulties arise from the nature of the problems themselves, and not from the method, and there is no apparent reason for this relation to be altered, so the particular advantages of the method should remain.

We use only such systems as are representable by linear differential equations, the difficulties of non linear systems being too great for elementary development. Consequently all coefficients are free of the time variable, and the physical constants of the system do not change. It is entirely possible that certain classes of non linear systems of almost periodic character can be treated by means of almost periodic functions, but they will not be considered at present.

The method seems particularly adapted to systems of an infinite number of degrees of freedom. Almost periodic systems of a finite number of degrees of freedom have a finite basis, and are well treated by other means, so will not be taken up for the time being.

In general the physical problem must be of finite extent, for in systems not so constituted, all the elements might be well behaved, and yet the energy be infinite. This energy consideration is of prime importance in the following method of almost periodic functions, and gives a grasp on the problem. Of course infinite systems can be so set up as to have finite energy, and these can be handled. When the energy is finite in a finite system, the displacements and derivatives must also be so, except possibly at a set of points of measure zero, and the con-

verse is also true. This gives us the key in the following method. In the infinite system, finite displacements and derivatives do not at all necessitate finite energy, and as will be seen, a necessary condition for the existence of an almost periodic system is the existence of a finite integral invariant of the energy type. Thus throughout the work dissipative forces are assumed absent, and there can be no friction, resistance, radiation, or other energy loss.

It is also necessary to note that for physical problems we are not dealing with almost periodic functions of the pure type. Periodicity in any case rests only in the time variable, while the space variables appear as parameters. This compels the replacing of ordinary differences by differences in the mean, and consequently the functions will be called functions almost periodic in the mean to distinguish from the regular almost periodic functions. The theory of functions almost periodic in the mean is worked out in Chapter I, up to the proof of the existence of a polynomial of trigonometric terms which approximates a given function. This is, of course, a direct analogue of Bohr's approximation formula for almost periodic functions, and is of fundamental importance in the application of functions almost periodic in the mean to physical problems.

The chief assumptions are those of finiteness of total energy of the system, and of the existence of such derivatives as may be necessary. Such assumptions will be seen to be not entirely arbitrary, for they are intimately connected with the permanence and possibility of existence of the system.

With linearity, finite energy, and permanence, we then show the existence of a finite time in which the system will return to within ϵ of its original condition; and then by showing the energy of the difference of these two positions to be small, we prove the displacement must remain small for successive vibrations. These assure the almost periodicity in the mean of the system, and the applicability of the method.

The latter parts of the method show the existence of the

solution of the system under the given boundary conditions, and result in the proof of the closure of the eigenfunktionen or characteristic functions.

The difficulty in the method lies almost entirely in the boundary problem, and this is so well appreciated, and so much met with in all branches of mathematical physics that it need not be enlarged upon. It is not a difficulty unique to any particular mode of attack upon physical problems, but arising in the very nature of the problem itself, must invariably be met, whatever the avenue of approach.

The author gratefully acknowledges the guidance and assistance of Professor Norbert Wiener in the preparation of this thesis.

Chapter I. Theory of Functions almost Periodic in the Mean

As has been pointed out in the introduction, the functions which occur generally in physical problems are not purely almost periodic, a given $y = f(t, x)$ being uniformly continuous in t, but continuous in the mean in x.

We will find it advantageous to assume all our functions to be real, but capable of assuming complex values, since we can write

$$f(t, x) = R(t, x) + iI(t, x).$$

The advantage of this conception is to allow the use of exponentials in place of trigonometric functions. Usually only R will appear in physical problems.

The displacement vectors are all parallel to the t axis, and are defined as follows. A displacement vector τ of $f(t, x)$, belonging to ϵ is a number such that

$$\int_{x_0}^{x_1} | f(t + \tau, x) - f(t, x) |^2 \, dx \leq \epsilon$$

for all values of t, and for all values of x between x_0 and x_1. A function of two variables x and t, for values as above stated,

will be called almost periodic in the mean, if for every ϵ there is a number r, such that in every region bounded by x_0 and x_1, and two lines parallel to the x axis and of distance apart r, such region being placed anywhere desired along the t axis, there is at least one displacement vector τ with initial point on the x axis, and terminus within this region. Such a region placed symmetrically about the origin is called the fundamental region. Placed elsewhere it is called an outer region. The distance r is called a displacement distance.

The following theorems will be proved for a function in two variables, t and x, but they are valid for t and any number of space parameters, as can readily be seen by noting that only the displacement region is changed in space dimensions by the introduction of more such parameters.

THEOREM I. *Every function almost periodic in the mean is bounded in the mean.*

We wish to show that

$$\int_{x_0}^{x_1} |f(t, x)|^2 \, dx \leq B.$$

We can use a finite region along the t axis, since this can at once be extended by means of our displacement vectors. The functions are assumed quadratically summable in x for every t. We know that

$$\int_{x_0}^{x_1} |f(t + \tau, x) - f(t, x)|^2 \, dx \leq \epsilon.$$

In a finite length r we know $f(t, x)$ is bounded by definition. Say

$$\int_{x_0}^{x_1} |f(t, x)|^2 \, dx \leq A.$$

Then by an inequality of the type $(a \pm b)^2 \leq 2a^2 + 2b^2$, we have from the original equation

$$\int_{x_0}^{x_1} |f(t + \tau, x)|^2 \, dx \leq 2A + 2\epsilon.$$

Now by means of translation vectors we can bring any other t inside the r region, so therefore

$$\int_{x_0}^{x_1} |f(t, x)|^2 \, dx \leq B \qquad \text{and} \qquad f(t, x)$$

is everywhere bounded in the mean.

THEOREM II. *Every function almost periodic in the mean is uniformly continuous in the mean.*

This conception of uniform continuity in the mean is derived from the usual definition of continuity in the mean, and an analogy to ordinary uniform continuity. The Borel theorem is instructive in the following.

Let r be the displacement distance for $\epsilon/9$ for $f(t, x)$. Let δ be less than r, and such that for any two strips of width dt, parallel to the x axis, and included between x_0 and x_1, within the double fundamental region $(t = \pm r; x = x_0, x_1)$, whose distance apart is less than δ, the function $f(t, x)$ has values differing by less than $\epsilon/9$. This is possible since $f(t, x)$ is continuous in t, and in the mean in x in this closed region. We may thus say that for this region the function is uniformly continuous in t, and uniformly continuous in the mean in x (see Borel theorem, as above). Now take two vertical strips of width dt as before, but outside the initial region. Their distance is again less than δ. We can find a displacement vector to take the strip at say t' inside the initial region to say a'. The same vector then takes the second strip at say T' inside the region to A', the distance between a' and A' being less than δ. Then we have

$$\int_{x_0}^{x_1} |f(a', x) - f(A', x)|^2 \, dx \leq \epsilon/9,$$

$$\int_{x_0}^{x_1} |f(t', x) - f(a', x)|^2 \, dx \leq \epsilon/9,$$

$$\int_{x_0}^{x_1} |f(A', x) - f(T', x)|^2 \, dx \leq \epsilon/9.$$

Now

$$\int (a + b + c)^2 \leq 3 \int (a^2 + b^2 + c^2).$$

Call the first bracket above a, the second b, and the third c. This gives at once

$$\int_{x_0}^{x_1} |\, a + b + c \,|^2\, dx = \int_{x_0}^{x_1} |\, f(t', x) - f(T', x) \,|^2\, dx \leq \epsilon$$

which proves the theorem.

COROLLARY. If τ_1 is a displacement vector for $f(t, x)$ pertaining to ϵ_1, there exists a positive number δ depending on ϵ_1 and $\epsilon_2 > \epsilon_1$ such that if $|\, \tau_2 - \tau_1 \,| < \delta$, then τ_2 is a displacement vector of ϵ_2.

For let δ be such that for any two strips at distance less than $\sqrt{2}\delta$ the function differs in the mean by less than $\epsilon_2 - \epsilon_1$. Then

$$\int_{x_0}^{x_1} |\, f(t + \tau_2, x) - f(t, x) \,|^2\, dx = \int_{x_0}^{x_1} |\, f(t + \tau_2, x) - f(t + \tau_1, x) + f(t + \tau_1, x) - f(t, x) \,|^2\, dx$$

$$\leq 2 \int_{x}^{x_1} |\, f(t + \tau_2, x) - f(t + \tau_1, x) \,|^2\, dx + 2 \int_{x_0}^{x_1} |\, f(t + \tau_1, x) - f(t, x) \,|^2\, dx \leq 2(\epsilon_2 - \epsilon_1) + 2\epsilon_1 = 2\epsilon_2.$$

THEOREM III. *Any two functions almost periodic in the mean are simultaneously almost periodic in the mean, that is, we may find an r such that in every displacement region there is a single displacement vector τ belonging to ϵ for both functions.*

Let δ be selected according to the Corollary to Theorem II, so that if τ_1 is a displacement vector for the first function $f_1(t, x)$ belonging to $\epsilon/2$, and $|\, \tau_2 - \tau_1 \,| < \delta$, then τ_2 will be a displacement vector for ϵ. Let r_1 and r_2 be displacement distances belonging to $\epsilon/4$ for $f_1(t, x)$ and $f_2(t, x)$ respectively. Construct the outer regions formed by drawing parallels to the x axis through the points $t = r, 2r, 3r, \cdots$. In each of these regions Q_j we may find a vector τ_1^j belonging to $\epsilon/4$ for $f_1(t, x)$, and a vector τ_2^j belonging to $\epsilon/4$ for $f_2(t, x)$. Now consider the set of vectors $T^j = \tau_2^j - \tau_1^j$, all drawn so that they have one extremity on the x axis. Their other extremities all lie within the double fundamental region $(t = \pm r;\ x = x_0,\ x_1)$. Draw

lines parallel to the x axis to divide this region into N rectangles q_i, each of length $x_1 - x_0$, and breadth $2r/N < \delta$. Some of these q_i will not contain any of the points T^i. For each of the q_i which contains any of these points, pick out the large rectangle Q_k nearest the origin with its T^k in q_i. Now take $A > r$ so large that a rectangle with the x axis for its middle line, and of breadth $2A$ contains all of the Q_k selected. We shall show that in every rectangle between x_0, x_1, and of breadth $4A$ there is at least one displacement vector belonging to ϵ for both functions.

Consider such a rectangle. In the rectangle of side $2A$ and the same middle line there is contained one of the rectangles Q_j. Find the T^j for this rectangle, and note the q_i in which it lies. There is a Q_k with its T^k in this same q_i. We may take $\tau = \tau_2{}^j - \tau_2{}^k$ as the common displacement vector sought. It is clearly inside the rectangle of breadth $4A$, since $\tau_2{}^j$ is within A of its middle line, while $\tau_2{}^k$ is numerically less than A. Since it is the difference of two displacement vectors for $f_2(t, x)$ belonging to $\epsilon/4$, it is also a displacement vector to $\epsilon/2$, and hence to ϵ for $f_2(t, x)$. Finally the vector $\tau' = \tau_1{}^j - \tau_1{}^k$ being the difference of two displacement vectors for $f_1(t, x)$ belonging to $\epsilon/4$ is a displacement vector for $f_1(t, x)$ belonging to $\epsilon/2$. But we have

$$| \tau - \tau' | = | \tau_2{}^j - \tau_2{}^k - \tau_1{}^j + \tau_1{}^k | = | T^j - T^k | < \delta.$$

Consequently from our choice of δ, the vector τ is a displacement vector for $f_1(t, x)$ belonging to ϵ.

COROLLARY. Any n functions almost periodic in the mean are simultaneously almost periodic in the mean, that is we may find an r such that in any region constructed from it there is a single displacement vector τ which belongs to ϵ for all n functions.

THEOREM IV. *The sum of N functions almost periodic in the mean is almost periodic in the mean.*

Let r be the displacement distance belonging to δ holding simultaneously for all n functions, u_1, u_2, $\cdots u_n$. τ is the common displacement vector belonging to ϵ/n^2. Both r and τ exist by the Corollary to Theorem III.

Then r is also the displacement distance for $\sum_{k=1}^{n} u_k$ considered as a function of t and x, for at two strips of width dt, parallel to the x axis, and differing by τ, each u_k differs in the mean by less than ϵ/n^2, so we have

$$\int_{x_0}^{x_1} |\sum u_k' - \sum u_k|^2 \, dx = \int_{x_0}^{x_1} |u_1' - u_1 + u_2' - u_2 + \cdots + u_n' - u_n|^2 \, dx.$$

Using an inequality of the type

$$\int (a_1 + a_2 + \cdots + a_n)^2 \leq n \int (a_1^2 + a_2^2 + \cdots + a_n^2)$$

$$\int_{x_0}^{x_1} |\sum u_k' - \sum u_k|^2 \, dx \leq n \int_{x_0}^{x_1} |u_1' - u_1|^2 \, dx + n \int_{x_0}^{x_1} |u_2' - u_2|^2 \, dx + n \int_{x_0}^{x_1} |u_3' - u_3|^2 \, dx + \cdots + n \int_{x_0}^{x_1} |u_n' - u_n|^2 \, dx \leq n \cdot n \cdot \epsilon/n^2 = \epsilon.$$

This proves the theorem

THEOREM V. *The product of a function almost periodic in the mean with a pure almost periodic function of one variable is almost periodic in the mean.*

The pure almost periodic function is defined only in the t, y plane, that is $y = f_1(t)$. Hence the restriction of the function almost periodic in the mean between x_0 and x_1 has no effect on the $f_1(t)$. The almost periodicity of each function resting alone in t, we can again by the Corollary to Theorem III find a common r, and a common τ belonging to ϵ for the two simultaneously. Then r is the displacement distance for the product $f_1(t)f_2(t, x)$ considered as a function of t and x. For as before take two strips of width dt, parallel to the x axis, and separated by the distance τ.

The integral we must examine is

$$\int_{x_0}^{x_1} |f_1(t+\tau)f_2(t+\tau, x) - f_1(t)f_2(t, x)|^2\,dx.$$

We know

$$|f_1(t+\tau) - f_1(t)| \leq \epsilon.$$

Subtracting and adding $f_1(t)f_2(t+\tau, x)$, the integral becomes

$$\begin{aligned}\int_{x_0}^{x_1} &|f_1(t+\tau)f_2(t+\tau, x) - f_1(t)f_2(t+\tau, x) \\ &+ f_1(t)f_2(t+\tau, x) - f_1(t)f_2(t, x)|^2\,dx \\ &= \int_{x_0}^{x_1} |f_2(t+\tau, x)[f_1(t+\tau) - f_1(t)] - f_1(t)[f_2(t+\tau, x) \\ &- f_2(t, x)]|^2\,dx \leq 2\int_{x_0}^{x_1} |f_2(t+\tau, x)\cdot\epsilon|^2\,dx \\ &+ 2\int_{x_0}^{x_1} |f_1(t)[f_2(t+\tau, x) - f_2(t, x)]|^2\,dx.\end{aligned}$$

The first integral on the right becomes vanishingly small with ϵ since f_2 is bounded in the mean. The second integral we attack through the boundedness of $f_1(t)$. Let $\max f_1(t) = A$. Then

$$\begin{aligned}2\int_{x_0}^{x_1} &|f_1(t)[f_2(t+\tau, x) - f_2(t, x)]|^2\,dx \\ &\leq 2A\int_{x_0}^{x_1} |f_2(t+\tau, x) - f_2(t, x)|^2\,dx \leq 2A\epsilon.\end{aligned}$$

Therefore we have

$$\int_{x_0}^{x_1} |f_1(t+\tau)f_2(t+\tau, x) - f_1(t)f_2(t, x)|^2\,dx \leq 2B\epsilon + 2A\epsilon,$$

where B is the upper mean bound of $f_2(t, x)$. This completes the proof.

NOTE. The product of two functions almost periodic in the mean will not in general be almost periodic in the mean. This is easily seen from the foregoing proof, since integrals of the type

$$\int_{x_0}^{x_1} |f_2(t+\tau, x)[f_1(t+\tau, x) - f_1(t, x)]|^2 dx$$

will not be vanishingly small unless $f_2(t+\tau, x)$ is bounded in the ordinary sense for all t, and all x between x_0 and x_1.

COROLLARY. The product of a function almost periodic in the mean with n pure almost periodic functions successively is almost periodic in the mean.

For the theory of pure almost periodic functions shows the product of any n almost periodic functions to be itself almost periodic.

COROLLARY. The absolute value of any function almost periodic in the mean is likewise almost periodic in the mean.

This follows from the same type of argument as used in Theorems IV and V, for the same r and τ will hold for the absolute value as for the given function.

THEOREM VI. *Any uniformly convergent series of functions almost periodic in the mean is almost periodic in the mean.*

Take N so large that the remainder after N terms in the series is less than $\epsilon/9$. By Theorem IV the sum S_N of the first N terms is a function almost periodic in the mean. So we may find a displacement distance r for it belonging to $\epsilon/9$. This will then be a displacement distance for S, the sum of the series to infinity, belonging to ϵ. For if τ is the displacement vector for S_N belonging to $\epsilon/9$ we have

$$\int_{x_0}^{x_1} |S_N(t+\tau, x) - S_N(t, x)|^2 dx \leq \epsilon/9.$$

But

$$\int_{x_0}^{x_1} |S(t, x) - S_N(t, x)|^2 dx \leq \epsilon/9,$$

and

$$\int_{x_0}^{x_1} |S(t+\tau, x) - S_N(t+\tau, x)|^2 dx \leq \epsilon/9.$$

Therefore

$$\int_{x_0}^{x_1} |S(t+\tau, x) - S(t, x)|^2 dx \leq \epsilon.$$

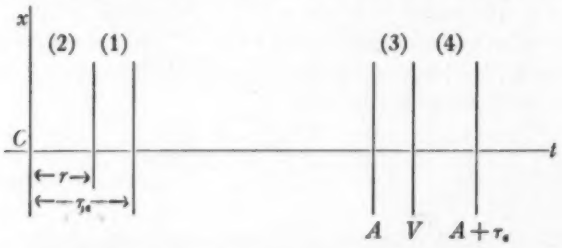
COROLLARY. A uniformly convergent series of continuous functions (continuous in a region including the upper mean bounds of the arguments) of functions almost periodic in the mean is almost periodic in the mean.

COROLLARY. If the series $\sum_{k,\,k} F_k(x) A_k e^{i\lambda_k t}$ converges uniformly for all values of x between x_0 and x_1, and hence in particular if $\sum_{k,\,k} |A_k F_k(x)|$ converges for the same values, it represents a function almost periodic in the mean.

THEOREM VII. *If $f(t, x)$ is almost periodic in the mean, then* $\lim_{T\to\infty} \frac{1}{T}\int_a^{a+T} f(t, x)dt$ *exists. We call if the mean value of $f(t, x)$ and abbreviate if by $M\{f(t, x)\}$. It is to be noticed that it is a function of x alone. Also the ordinary mean value as given by* $\lim_{T\to\infty} \frac{1}{T}\int_0^{T} f(t, x)dt$ *exists, and we have*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^{T} f(t, x)dt = \lim_{T\to\infty} \frac{1}{T}\int_a^{a+T} f(t, x)dt.$$

Let r be the displacement distance pertaining to ϵ for $f(t, x)$. Let the first displacement vector that is greater than r be τ_ϵ.



Take any region along t, starting say at A, and ending at $A + \tau_\epsilon$, as shown. We will now consider the integral

$$\int_{x_0}^{x_1} \left| \frac{1}{\tau_\epsilon}\int_A^{A+\tau_\epsilon} f(t, x)dt - \frac{1}{\tau_\epsilon}\int_0^{\tau_\epsilon} f(t, x)dt \right|^2 dx.$$

Let V be a displacement vector carrying some interior point

of the A, $A + \tau_\epsilon$ region back to the origin. We now split the above integral up as follows

$$\int_{x_0}^{x_1} \left| \frac{1}{\tau_\epsilon} \int_V^{A+\tau_\epsilon} f(t, x)dt - \frac{1}{\tau_\epsilon} \int_0^{A+\tau_\epsilon-V} f(t, x)dt \right.$$
$$(4) \qquad (2)$$
$$\left. + \frac{1}{\tau_\epsilon} \int_A^V f(t, x)dt - \frac{1}{\tau_\epsilon} \int_{A+\tau_\epsilon-V}^{\tau_\epsilon} f(t, x)dt \right|^2 dx.$$
$$(3) \qquad (1)$$

Small numbers, corresponding to the region of integration marked on the figure, are placed under each of the above integrals for clarity. It is evident that the region at O, and that at A have each been divided into two parts.

Since $\tau_\epsilon > r$, its existence is assured, and since the outer region at A also has a length τ_ϵ, then V exists. From our fundamental definition of a function almost periodic in the mean the points in A, $A + \tau_\epsilon$ can be translated into O, $O + \tau_\epsilon$ with the mean squared difference between the two positions less than ϵ, or 4ϵ, depending on whether we must employ one translation or two in order to move the points back. The above integral can then be written

$$\int_{x_0}^{x_1} \left| \frac{1}{\tau_\epsilon} \int_0^{A+\tau_\epsilon-V} [f(t + V, x) - f(t, x)]dt \right.$$
$$\left. + \frac{1}{\tau_\epsilon} \int_{A+\tau_\epsilon-V}^{\tau_\epsilon} [f(t + V - \tau_\epsilon, x) - f(t, x)]dt \right|^2 dx$$
$$\leq \frac{2}{\tau_\epsilon^2} \int_{x_0}^{x_1} \left| \int_0^{A+\tau_\epsilon-V} [f(t + V, x) - f(t, x)]dt \right|^2 dx.$$

Now since $\int_0^{A+\tau_\epsilon-V} dt \leq \tau_\epsilon$, using the Schwarz inequality on this last integral above reduces it to

$$\leq 2/\tau_\epsilon \int_{x_0}^{x_1} dx \int_0^{A+\tau_\epsilon-V} |f(t + V, x) - f(t, x)|^2 dt$$
$$\leq \frac{2b}{\tau_\epsilon} \int_0^{A+\tau_\epsilon-V} (4\epsilon)dt \leq 8b\epsilon,$$

where b is the breadth of the region considered, along the x axis. Therefore we have proved

$$\int_{x_0}^{x_1} \left| \frac{1}{\tau_\epsilon} \int_A^{A+\tau_\epsilon} f(t, x)dt - \frac{1}{\tau_\epsilon} \int_0^{\tau_\epsilon} f(t, x)dt \right|^2 dx \leq \epsilon_1,$$

for an arbitrarily chosen region along t.

If now we take any length of time desired, we may divide it into an integral number of intervals of length τ_ϵ, for all of which the above relation holds, and a remainder interval. For the first part then we know that

$$\int_{x_0}^{x_1} \left| \frac{1}{n\tau_\epsilon} \int_A^{A+n\tau_\epsilon} f(t, x)dt - \frac{1}{\tau_\epsilon} \int_0^{\tau_\epsilon} f(t, x)dt \right|^2 dx \leq \epsilon_2$$

by an exactly similar proof to the above, for the points in the interval A, $A + n\tau_\epsilon$ can be brought back to the interval O, $O + \tau_\epsilon$ such that the mean squared difference of the positions can be made as small as desired by properly choosing r, and hence ϵ.

Now consider the remainder piece. Call its length $\theta\tau_\epsilon$, where $O < \theta < 1$. Using the same attack as before we set up

$$\int_{x_0}^{x_1} \left| \frac{1}{n\tau_\epsilon} \int_A^{A+n\tau_\epsilon} f(t, x)dt - \frac{1}{n\tau_\epsilon + \theta\tau_\epsilon} \int_A^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t, x)dt \right|^2 dx.$$

The latter integral we are not prepared to handle, so we will change both integrals, though still covering the same region. Let the first become

$$\frac{1}{n\tau_\epsilon} \int_{A+\theta\tau_\epsilon}^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t, x)dt.$$

Then the second changes to

$$\frac{1}{n\tau_\epsilon + \theta\tau_\epsilon} \int_A^{A+n\tau_\epsilon} f(t, x)dt - \frac{1}{n\tau_\epsilon} \int_A^{A+n\tau_\epsilon} f(t, x)dt$$
$$= -\frac{\theta\tau_\epsilon}{n\tau_\epsilon(n\tau_\epsilon + \theta\tau_\epsilon)} \int_A^{A+n\tau_\epsilon} f(t, x)dt.$$

Placing these two expressions in the original one, and using the inequality $\frac{1}{2}(a + b)^2 \leq a^2 + b^2$ we get

$$\int_{x_0}^{x_1}\left|\frac{1}{n\tau_\epsilon}\int_A^{A+n\tau_\epsilon} f(t,x)dt - \frac{1}{n\tau_\epsilon+\theta\tau_\epsilon}\int_A^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t,x)dt\right|^2 dx$$
$$\leq 2\int_{x_0}^{x_1}\left|\frac{1}{n\tau_\epsilon}\int_{A+\theta\tau_\epsilon}^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t,x)dt\right|^2 dx$$
$$+ 2\int_{x_0}^{x_1}\left|\frac{\theta\tau_\epsilon}{n\tau_\epsilon(n\tau_\epsilon+\theta\tau_\epsilon)}\int_A^{A+n\tau_\epsilon} f(t,x)dt\right| dx.$$

The interval in each case now being $n\tau_\epsilon$, each integral on the right is less than or equal to

$$\frac{1}{n}\max\int_{x_0}^{x_1}|f(t,x)|^2\,dx$$

and the whole original expression is less than or equal to

$$\frac{4}{n}\max\int_{x_0}^{x_1}|f(t,x)|^2\,dx.$$

Since $f(t,x)$ is bounded in the mean, the value here approaches zero as n approaches infinity. Thus it follows at once

$$\underset{B-A\to\infty}{\text{l.i.m.}}\int_{x_0}^{x_1}\left|\frac{1}{\tau_\epsilon}\int_\lambda^{\tau_\epsilon} f(t,x)dt - \frac{1}{B-A}\int_A^B f(t,x)dt\right|^2 dx \leq \lambda\epsilon.$$

Therefore

$$\underset{B-A\to\infty}{\text{l.i.m.}}\int_A^B f(t,x)dt \quad \text{exists.}$$

Also, since we have used the integral of form

$$\frac{1}{\tau_\epsilon}\int_A^{A+\tau_\epsilon} f(t,x)dt$$

throughout in our differences in the mean, the sharper mean value is established. That is we have proved the existence of

$$\underset{\alpha\to\infty}{\text{l.i.m.}}\frac{1}{\alpha}\int_A^{A+\alpha} f(t,x)dt.$$

We must now show that any function almost periodic in the mean can be expressed by means of a trigonometric series. The

method is that of N. Wiener.[7] The first parts of the method can be taken over at once as they stand. We see that the whole energy of the spectrum of such a function is contained in a denumerable set of lines, this idea of spectrum being the one given in the paper mentioned. Finally we employ the Parseval theorem to prove the existence of the trigonometric polynomial approximation.

Starting with a given function almost periodic in the mean, $f(t, x)$, we form the function

$$\phi(t) = \max_z \int_{x_0}^{x_1} |f(t+z, x) - f(z, x)|^2 \, dx.$$

Employing the translation number τ of $f(t, x)$

$$|\phi(t+\tau) - \phi(t)| \leq 2 \max_z \int_{x_0}^{x_1} \{|f(t+z+\tau, x) - f(z+\tau, x)| - |f(t+z, x) - f(z, x)|\} \cdot \{|f(t+z+\tau, x) - f(z+\tau, x)| + |f(t+z, x) - f(z, x)|\} dx.$$

Using the Schwarz inequality this becomes

$$|\phi(t+\tau) - \phi(t)| \leq 2 \sqrt{\max_z \int_{x_0}^{x_1} \{|f(z+t+\tau, x) - f(z+\tau, x)| - |f(t+z, x) - f(z, x)|\}^2 dx \cdot \max_z \int_{x_0}^{x_1} \{|f(z+t+\tau, x) - f(z+\tau, x)| + |f(t+z, x) - f(z, x)|\}^2 dx}.$$

To evaluate this let us consider each integral under the radical separately. The first is

$$\max_z \int_{x_0}^{x_1} \{|f(t+z+\tau, x) - f(z+\tau, x)| - |f(t+z, x) - f(z, x)|\}^2 dx \leq \max_z \int_{x_0}^{x_1} \{|f(t+z+\tau, x) - f(z+\tau, x)| + |f(t+z, x) - f(z, x)|\}^2 dx \leq 4 \max_z \int_{x_0}^{x_1} |f(z+\tau, x) - f(z, x)|^2 dx.$$

The second integral may be handled as it stands, for

$$\max_z \int_{x_0}^{x_1} \{|f(t+z+\tau, x) - f(z+\tau, x)| + |f(t+z, x) - f(z, x)|\}^2 dx \leq 4 \max_z \int_{x_0}^{x_1} |f(z+\tau, x) - f(z, x)|^2 dx \cdot \leq 4 \max_z \int_{x_0}^{x_1} |f(z, x)|^2 ax.$$

Therefore we may now write the original expression as

$$|\phi(t+\tau) - \phi(t)| \leq 2\sqrt{4 \max_z \int_{x_0}^{x_1} |f(z+\tau, x) - f(z, x)|^2 dx \cdot 4 \max_z \int_{x_0}^{x_1} |f(z, x)|^2 dx}.$$

But since τ is a translation number for $f(t, x)$

$$4 \max_z \int_{x_0}^{x_1} |f(z+\tau, x) - f(z, x)|^2 dx \leq 4\epsilon.$$

Since $f(t, x)$ is bounded in the mean we may write

$$4 \max_z \int_{x_0}^{x_1} |f(z, t)|^2 dx = \frac{A^2}{16} = \text{const.}$$

Therefore

$$|\phi(t+\tau) - \phi(t)| \leq 2\sqrt{4\epsilon \cdot A^2/16} = A\sqrt{\epsilon}$$

and we see that any translation number τ of $f(t, x)$ pertaining to ϵ, is also a translation number for $\phi(t)$ pertaining to $\sqrt{\epsilon}$ multiplied by a constant which depends only on the function $f(t, x)$.

Now any polynomial in functions almost periodic in the mean is almost periodic in the mean. Moreover the uniform limit of such functions is almost periodic in the mean, and any continuous function of them is almost periodic in the mean.

Let $H(u)$ be a continuous, non negative function of u which is 1 when $O \leq u \leq \epsilon/2$, between 1 and 0 when $\epsilon/2 \leq u \leq \epsilon$,

and 0 for all $u > \epsilon$. From our previous function $\phi(t)$ set up

$$\psi_\epsilon(t) = \frac{H[\phi(u)]}{MH[\phi(u)]},$$

where M signifies the mean value.

Since $H[\phi(t)]$ is somewhere positive, and everywhere non negative, $MH[\phi(t)]$ can not vanish. Then from the discussion above ψ_ϵ is almost periodic. Now form the function

$$f_\epsilon(t, x) = M[f(z, x)\psi_\epsilon(t - z)],$$

where as before z is a time variable. The existence and almost periodicity in the mean of $f_\epsilon(t, x)$ are also easily established as above. The definition of ψ_ϵ ensures that

$$\int_{x_0}^{x_1} |f(t, x) - f(z, x)|^2 dx \leq \mathrm{X}(\epsilon)$$

provided $\psi_\epsilon(t - z) \neq 0$. Here $\mathrm{X}(\epsilon) \to 0$ as $\epsilon \to 0$. Now since from the definition of $f_\epsilon(t, x)$ we see that it is a mean of these values of $f(z, x)$ we have

$$\max_t \int_{x_0}^{x_1} |f(t, x) - f_\epsilon(t, x)|^2 dx \leq \mathrm{X}(\epsilon).$$

Similarly define the function

$$f^{(\epsilon)}(t, x) = M[f_\epsilon(t, x)\psi_\epsilon(t - z)].$$

Then as before $f^{(\epsilon)}$ exists and is almost periodic in the mean. Then we have at once

$$\max_t \int_{x_0}^{x_1} |f^{(\epsilon)}(t, x) - f_\epsilon(t, x)|^2 dx \leq \mathrm{X}(\epsilon).$$

From this relation there results

$$\max_t \int_{x_0}^{x_1} |f^{(\epsilon)}(t, x) - f(t, x)|^2 dx \leq 4\mathrm{X}(\epsilon).$$

Now let us substitute the value of $f_\epsilon(t, x)$ into the equation for $f^{(\epsilon)}(t, x)$.

$$f^{(\epsilon)}(t, x) = \underset{z}{M}\left[\psi_\epsilon(t - z) \underset{\tau}{M} \{f(\tau, x)\psi_\epsilon(z - \tau) \right].$$

Due to the uniformity properties of functions almost periodic in the mean we may interchange as follows

$$f^{(\epsilon)}(t, x) = \underset{\tau}{M}\left[f(\tau, x) \underset{z}{M} \{\psi_\epsilon(t - z)\psi_\epsilon(z - \tau)\} \right].$$

Referring to part 4 of Professor Wiener's paper we see that

$$\underset{z}{M} \{\psi_\epsilon(t - z)\psi_\epsilon(z - \tau)\} = \sum_0^\infty a_k \cos \lambda_k(t - \tau).$$

where all the coefficients are positive, and $\sum_0^\infty a_k$ converges. So

$$f^{(\epsilon)}(t, x) = \underset{\tau}{M}\left[f(\tau, x) \sum_0^\infty a_k \cos \lambda_k(t - \tau) \right].$$

Now since we know that

$$\left| \frac{1}{2T} \int_{-T}^{T} f(\tau, x) \cos \lambda_k(t - \tau) dt \right| \leq \max |f(\tau, x)|,$$

we can invert the order of M and $\sum$. This gives

$$f^{(\epsilon)}(t, x) = \sum_0^\infty a_k \{\cos \lambda_k t M[f(\tau, x) \cos \lambda_k \tau] + \sin \lambda_k t M[f(\tau, x) \sin \lambda_k \tau]\},$$

and inasmuch as

$$|M[f(\tau, x)^{\sin}_{\cos} \lambda_k \tau]| \leq |f(\tau, x)|,$$

then $f(t, x)$ is the sum of an absolutely convergent series of trigonometric terms. That is we can choose N so large that

$$\max_t \int_{x_0}^{x_1} \left| f^{(\epsilon)}(t, x) - \sum_1^N a_k \{\cos \lambda_k t M[f(\tau, x) \cos \lambda_k \tau] + \sin \lambda_k t M[f(\tau, x) \sin \lambda_k \tau]\} \right|^2 dx \leq \mathrm{X}(\epsilon)$$

Therefore, using the equation previously obtained

$$\max_t \int_{x_0}^{x_1} |f^{(\epsilon)}(t, x) - f(t, x)|^2 dx \leq 4X(\epsilon).$$

We get the final result

$$\max_t \int_{x_0}^{x_1} \Big| f(t, x) - \sum_1^N a_k \{\cos \lambda_k t M[f(\tau, x) \cos \lambda_k \tau] + \sin \lambda_k t M[f(\tau, x) \sin \lambda_k \tau]\} \Big|^2 dx \leq 10X(\epsilon).$$

That is, given an ϵ, there is a polynomial

$$P_\epsilon(t, x) = \sum_1^N A_k(x) e^{i\Lambda_k t},$$

such that

$$\int_{x_0}^{x_1} |f(t, x) - P_\epsilon(t, x)|^2 dx \leq \epsilon.$$

This proves Bohr's approximation theorem for a function almost periodic in the mean, and shows we may approximate uniformly to any desired degree of accuracy to a $f(t, x)$ by means of a trigonometric polynomial.

Chapter II. Proof of the almost Periodicity in the Mean of a Loaded Vibrating String

Let us now consider a vibrating string whose mass and tension are functions of the length, but not of the time. Call the mass function $\mu(x)$, and the tension $T(x)$. Then from $F = ma$ we have

$$\frac{\partial}{\partial x}\left[T(x)\frac{\partial y}{\partial x}\right] = \mu(x)\frac{\partial^2 y}{\partial t^2}. \tag{1}$$

For stability $\mu(x)$ and $T(x)$ are always finite and greater than 0. Then the kinetic energy of the system is, for distance x_0, x_1 between the ends

$$\frac{1}{2}\int_{x_0}^{x_1} \mu(x)\left(\frac{\partial y}{\partial t}\right)^2 dx. \tag{2}$$

Similarly the potential energy is

$$\frac{1}{2}\int_{x_0}^{x_1} T(x)\left(\frac{\partial y}{\partial x}\right)^2 dx. \tag{3}$$

Thus the total energy, assumed finite, is

$$E_0 = \frac{1}{2}\int_{x_0}^{x_1}\left[\mu(x)\left(\frac{\partial y}{\partial t}\right)^2 + T(x)\left(\frac{\partial y}{\partial x}\right)^2\right] dx, \tag{4}$$

Knowing that

$$\frac{\partial}{\partial t}[\mu(x)] = 0, \qquad \text{and} \qquad \frac{\partial}{\partial t}[T(x)] = 0$$

$$\frac{\partial E_0}{\partial t} = \int_{x_0}^{x_1}\left[\mu(x)\left(\frac{\partial y}{\partial t}\right)\left(\frac{\partial^2 y}{\partial t^2}\right) + T(x)\left(\frac{\partial y}{\partial x}\right)\left(\frac{\partial^2 y}{\partial x \partial t}\right)\right] dx. \tag{5}$$

Substituting from (1)

$$\frac{\partial E_0}{dt} = \int_{x_0}^{x_1}\left\{\left(\frac{\partial y}{\partial t}\right)\frac{\partial}{\partial x}\left[T(x)\left(\frac{\partial y}{\partial x}\right)\right] + \left[T(x)\left(\frac{\partial y}{\partial x}\right)\right]\frac{\partial}{\partial x}\left(\frac{\partial y}{\partial t}\right)\right\} dx, \tag{6}$$

$$\frac{\partial E_0}{\partial t} = \int_{x_0}^{x_1}\frac{\partial}{\partial x}\left[T(x)\frac{\partial y}{\partial x}\frac{\partial y}{\partial t}\right] dx = \left[T(x)\frac{\partial y}{\partial x}\frac{\partial y}{\partial t}\right]_{x_0}^{x_1}. \tag{7}$$

Now at $x = x_0$, and $x = x_1$, $\partial y/\partial t = 0$, therefore $\partial E_0/\partial t = 0$ and the energy is invariant, as we could have predicted since there is no dissipation.

Following the general outline of the method just used we will now assume the first and second energies to be finite. Taking the first derivative of (1) with respect to time, and forming the energy equation we have

$$E_1 = \frac{1}{2}\int_{x_0}^{x_1}\left[\mu(x)\left(\frac{\partial^2 y}{\partial t^2}\right)^2 + T(x)\left(\frac{\partial^2 y}{\partial x \partial t}\right)^2\right] dx. \tag{8}$$

Similarly the second energy is

$$E_2 = \frac{1}{2}\int_{x_0}^{x_1}\left[\mu(x)\left(\frac{\partial^3 y}{\partial t^3}\right)^2 + T(x)\left(\frac{\partial^3 y}{\partial x \partial t^2}\right)^2\right] dx. \tag{9}$$

We now wish to show these invariant.

$$\frac{\partial E_1}{\partial t} = \int_{x_0}^{x_1} \left[\mu(x) \left(\frac{\partial^2 y}{\partial t^2} \right) \left(\frac{\partial^3 y}{\partial t^3} \right) + T(x) \left(\frac{\partial^2 y}{\partial x \partial t} \right) \left(\frac{\partial^3 y}{\partial x \partial t^2} \right) \right] dx. \quad (10)$$

The first two derivatives of (1) are

$$\frac{\partial}{\partial x} \left[T(x) \frac{\partial^2 y}{\partial x \partial t} \right] = \mu(x) \frac{\partial^3 y}{\partial t^3}, \quad (11)$$

and

$$\frac{\partial}{\partial x} \left[T(x) \frac{\partial^3 y}{\partial x \partial t^2} \right] = \mu(x) \frac{\partial^4 y}{\partial t^4}. \quad (12)$$

Using (11) in (10) we obtain

$$\begin{aligned} \frac{\partial E_1}{\partial t} &= \int_{x_0}^{x_1} \left\{ \left(\frac{\partial^2 y}{\partial t^2} \right) \frac{\partial}{\partial x} \left[T(x) \frac{\partial^2 y}{\partial x \partial t} \right] \right. \\ &\qquad \left. + \left[T(x) \left(\frac{\partial^2 y}{\partial x \partial t} \right) \right] \frac{\partial}{\partial x} \left(\frac{\partial^2 y}{\partial t^2} \right) \right\} dx, \\ &= \int_{x_0}^{x_1} \frac{\partial}{\partial x} \left[T(x) \left(\frac{\partial^2 y}{\partial x \partial t} \right) \left(\frac{\partial^2 y}{\partial t^2} \right) \right] dx, \\ &= \left[T(x) \left(\frac{\partial^2 y}{\partial x \partial t} \right) \left(\frac{\partial^2 y}{\partial t^2} \right) \right]_{x_0}^{x_1} = 0 \end{aligned} \quad (13)$$

as before since $\partial^2 y / \partial t^2 = 0$ when $x = x_0$ and $x = x_1$. The same will hold for the second energy and any higher which we may assume to exist and be finite, for we always have

$$\frac{\partial E_n}{\partial t} = \left[T(x) \left(\frac{\partial^{n+1} y}{\partial x \partial t^n} \right) \left(\frac{\partial^{n+1} y}{\partial t^{n+1}} \right) \right]_{x_0}^{x_1} = 0. \quad (14)$$

Therefore all the energies are invariant.

We now wish to show that the integral of the squares of certain elements of our system is bounded. The square is used since we must later work in the mean with these elements. We will assume the existence of various derivatives, which will be introduced as the need arises.

The energy E_0 is finite, and $\mu(x)$ and $T(x)$ are positive:

Hence

$$\frac{1}{2}\int_{x_0}^{x_1} \mu(x)\left(\frac{\partial y}{\partial t}\right)^2 dx \geq 0,$$

Then

$$\frac{1}{2}\int_{x_0}^{x_1} T(x)\left(\frac{\partial y}{\partial x}\right)^2 dx \leq E,$$

and calling the minimum value of $T(x)$ by T

$$\frac{1}{2}\int_{x_0}^{x_1} T\left(\frac{\partial y}{\partial x}\right)^2 dx \leq E.$$

This gives us the worst possible case, and we see that

$$\int_{x_0}^{x_1}\left(\frac{\partial y}{\partial x}\right)^2 dx \leq 2E/T. \tag{15}$$

Thus $\int_{x_0}^{x_1}\left(\frac{\partial y}{\partial x}\right)^2 dx$ is everywhere bounded in the interval x_0, x_1.

An exactly similar argument holds for $\int_{x_0}^{x_1}\left(\frac{\partial y}{\partial t}\right)^2 dx$, so it is similarly bounded. (15a)

To consider the boundedness of the ordinate we consider

$$y^2 = \left[\int_0^x \left(\frac{\partial y}{dx}\right)\partial x\right]^2. \tag{16}$$

By the Schwartz inequality

$$y^2 = \int_0^x dx \int_0^x \left(\frac{\partial y}{\partial x}\right)^2 dx \leq x\int_0^x \left(\frac{\partial y}{\partial x}\right)^2 dx. \tag{17}$$

Then for the interval $x = x_0, x_1$ we see from (15) that y^2, and thus y is bounded.

Now by employing E_1 and E_2 from (8) and (9), and by exactly similar procedure to that producing the results of (15) and (15a) we at once prove the boundedness of

$$\int\left(\frac{\partial^2 y}{\partial t^2}\right)^2 dx, \quad \int\left(\frac{\partial^3 y}{\partial t^3}\right)^2 dx, \quad \int\left(\frac{\partial^2 y}{\partial x \partial t}\right)^2 dx, \quad \text{and} \quad \int\left(\frac{\partial^3 y}{\partial x \partial t^2}\right)^2 dx.$$

These will be used in the next step of the development. We now prove

LEMMA

If $f(t, x)$ is developable in a Fourier's series, and if $\int \left[\frac{\partial f(t, x)}{\partial x}\right]^2 dx$ is finite, then the remainder of the series is less than a quantity depending on $(\partial f/\partial x)^2$, and on the number of terms of the series, and vanishes uniformly as the number of terms becomes infinite.

Proof: Let $f(t, x)$ be expanded in the series $\sum a_n(t)e^{i\Lambda_n x}$. Then differentiating:

$$\frac{\partial f}{\partial x} = \sum a_n(t) i \Lambda_n e^{i\Lambda_n x}.$$

Now $\int \left(\frac{\partial f}{\partial x}\right)^2 dx$ is bounded by hypothesis, and the a_0 term is bounded since it is an average. Therefore we may write that $\sum \Lambda_n^2 a_n^2(t) < A$, where the constant A depends only on $(\partial f/\partial x)^2$ since the maximum values of $ie^{i\Lambda_n x}$ are 1 and -1.

Therefore $\sum_{N}^{\infty} a_n^2(t) < A/N^2$, and $\sum_{-N}^{-\infty} a_n^2(t) < A/N^2$.

Consequently $\sum_{-N}^{N} [a_n(t)]^2$ is uniformly bounded, and the remainder is less than $2A/N^2$.

Q. E. D.

Equation (14) shows the invariance of any energy we wish to assume existant and finite. From this, by means of the procedure leading to results (15), etc., we have shown the boundedness of any expression of the type $\int \left[\frac{\partial}{\partial x}\left(\frac{\partial^n y}{\partial t^n}\right)\right]^2 dx$, or of the type $\int \left(\frac{\partial^n y}{\partial t^n}\right)^2 dx$. Our Lemma applies at once to all of the former kind. We are now in a position to prove that the system returns to its original position (within ϵ) in a finite time.

We will then later show that it so remains, or that the energy of the difference of the two positions after a finite time remains small.

We define as the distance between two functions

$$S = \int\Big[|y - y_1|^2 + \left|\frac{\partial y}{\partial x} - \frac{\partial y_1}{\partial x}\right|^2 + \left|\frac{\partial y}{\partial t} - \frac{\partial y_1}{\partial t}\right|^2 + \left|\frac{\partial^2 y}{\partial t^2} - \frac{\partial^2 y_1}{\partial t^2}\right|^2\Big] dx. \quad (18)$$

That is, by distance we mean the differences in the mean between the position, slope, velocity, and acceleration of the system.

We have shown that $\sum_{-N}^{N}[a_n(t)]^2$ is uniformly bounded for the ordinate series. We can now expand the other elements of the distance function (18) in similar series, and then by differentiating and using the lemma can show that their respective series are also bounded for $\sum_{-N}^{N}$.

Now for any of these we may consider the $(2N + 1)$ a's to be the coordinates of a point in hyperspace, and since they are bounded, as lying on or in a sphere in this space. Consider this sphere to be divided into compartments of dimensions ϵ, where lengths are defined by S in (18).

The definition of S shows each compartment to be a combination of the compartments of y, $\partial y/\partial x$, $\partial y/\partial t$, and $\partial^2 y/\partial t^2$ respectively, that is, with every value of one of the elements is associated a value of each of the others. Now give each compartment a number such that this number shows the time for the vector A to reach it, starting from $t = 0$. If then we take the largest time of all these, the vector will have been in all the spaces it will ever enter. So there is a finite time during which A will have approached every point of its trajectory closer than ϵ. Thus, given any ϵ, a finite time exists in which A will be closer than ϵ to any point it possibly can approach. That is,

all the elements will have returned simultaneously to within ϵ of their original values, in the mean. Consequently, from (4) the energy will also return to zero, for the difference of the two positions.

We must now show that once all the elements have returned to very near the original position, they will so remain during successive vibrations of the string. The method is to show that the energy of the difference of any one of the elements stays always small. We will use the ordinate y.

Take equation (1), divide through by $\mu(x)$, square both sides, and integrate with respect to x. Then we have

$$\int\left[\frac{1}{\mu(x)}\right]^2\left[\frac{\partial}{\partial x}\left\{T(x)\frac{\partial y}{\partial x}\right\}\right]^2 dx = \int\left[\frac{\partial^2 y}{\partial t^2}\right]dx. \tag{19}$$

We know from our previous discussion that

$$\int\left|\frac{\partial y_1}{\partial x}-\frac{\partial y}{\partial x}\right|^2 dx < \epsilon, \quad \text{and} \quad \int\left|\frac{\partial y_1}{\partial t}-\frac{\partial y}{\partial t}\right|^2 dx < \epsilon,$$

Then

$$\int [T(x)]^2\left|\frac{\partial}{\partial x}(y-y_1)\right|^2 dx < \phi_1(\epsilon), \tag{20}$$

and also

$$\int\left[\frac{1}{\mu(x)}\right]^2\left|\frac{\partial}{\partial x}\left\{T(x)\frac{\partial(y-y_1)}{\partial x}\right\}\right|^2 dx = \int\left|\frac{\partial^2(y-y_1)}{\partial t^2}\right|^2 dx < \phi_2(\epsilon) \tag{21}$$

where $\phi_1(\epsilon)$ and $\phi_2(\epsilon)$ approach 0 as ϵ approaches 0.

We now define the potential energy function

$$u(x) = T(x)\frac{\partial}{\partial x}(y-y_1). \tag{22}$$

Using this expression (21) becomes

$$\int\left[\frac{1}{\mu(x)}\right]^2\left|\frac{\partial u}{\partial x}\right|^2 dx < \phi_1(\epsilon) \tag{23}$$

We will now express $u(x)$ as

$$u(x) = \int_P^x \left(\frac{\partial u}{\partial x}\right) dx, \tag{24}$$

Then

$$\int_P^x |u(x)|^2 dx = \int_{x_0}^{x_1} \left| \int_P^x \left(\frac{\partial u}{\partial x}\right) dx \right|^2 dx. \tag{25}$$

Using the Schwartz inequality we obtain

$$\int |u(x)|^2 dx = \int_{x_0}^{x_1} dx \left| \int_{x_0}^{x_1} d\xi \int_{x_0}^{x} \left(\frac{\partial u}{\partial \xi}\right) d\xi \right|$$

$$\leq (x_1 - x_0)^2 \int_{x_0}^{x_1} \left| \left(\frac{\partial u}{\partial \xi}\right) \right|^2 d\xi. \tag{26}$$

Then from (23) we see that the potential energy is small for the differences of position of the system during successive motions separated by a translation vector. In the same way we can handle the kinetic energy, so that it is seen that the total energy of the differences of position of the system is small for all repeated motions.

Now considering our definition of a function almost periodic in the mean, we see that from the argument of the return of the elements of the distance between two positions of our system, that the largest time for the vector A to reach each point of its trajectory corresponds to the displacement distance r for the given dimension ϵ of the compartments. The fact that the energy of the difference of positions remains small for the succeeding motion of the system establishes the existence of the displacement vector τ.

That is, taking the largest time, r, as above, we know there is a time equal to it, or less than it, when all the elements of the initial position are simultaneously within ϵ of that position in the mean. Thus we have $\tau_1 \leq r$, and using this τ_1 we know that

$$\int_{x_0}^{x_1} |f(t + \tau_1, x) - f(t, x)|^2 dx \leq \epsilon.$$

If we wish, we may now start from this new position $f(t+\tau_1, x)$ and repeat the same argument. The largest time from this position will of course be r again, provided we keep the same ϵ. Thus there will be a new τ_2 such that

$$\int_{x_0}^{x_1} |f(t+\tau_2, x) - f(t+\tau_1, x)|^2 dx \leq \epsilon,$$

or carrying back to the original state

$$\int_{x_0}^{x_1} |f(t+\tau_2, x) - f(t, x)|^2 dx = \int_{x_0}^{x_1} |f(t+\tau_2, x) - f(t+\tau_1, x) + f(t+\tau_1, x) - f(t, x)|^2 dx \leq 2\int_{x_0}^{x_1} |f(t+\tau_2, x) - f(t+\tau_1, x)|^2 dx + 2\int_{x_0}^{x_1} |f(t+\tau_1, x) - f(t, x)|^2 dx \leq 4\epsilon.$$

Or using the fundamental definition of a function almost periodic in the mean we may go directly from the last state above to the initial one by letting A make the circuit of its path twice. This amounts to taking the region from r out away from the origin. Then we have a vector τ_2' in this region such that

$$\int_{x_0}^{x_1} |f(t+\tau_2', x) - f(t, x)|^2 dx \leq \epsilon.$$

Since this can evidently be carried out for all t, it proves the given system to be almost periodic in the mean.

Chapter III. Existence of the Solution of the Problem

We have shown that if a solution exists for the system under consideration it will be almost periodic in the mean. We will now consider the differential equation to prove this existence, for any initial position of the string desired, if only this position be often enough differentiable in x.

Our original equation is linear, and of the hyperbolic type. We know the solution for $t = 0$, for it reduces to the initial position. We also know the value of the ordinate, velocity, and

acceleration along the two lines $x = x_0, x_1$, for all values of t. Assuming the string to be at rest before $t = 0$, to have the values $y = f(x)$ at $t = 0$, and then to be released, the whole motion is included in the region $x = x_0, x_1, t = 0, \infty$.

In order to handle the equation better we will transform it by a rotation of the axes through 45 degrees, and a suitable distortion of the parametric direction. This will eliminate the $\partial^2 y/\partial t^2$ and $\partial^2 y/\partial x^2$, and introduce $\partial^2 y/\partial \xi \partial \eta$ instead.

Introducing the new variables ξ and η we get the auxiliary equations

$$\left.\begin{aligned} \mu(x)\left(\frac{\partial \xi}{\partial t}\right)^2 &= T(x)\left(\frac{\partial \xi}{\partial x}\right)^2 \\ \mu(x)\left(\frac{\partial \eta}{\partial t}\right)^2 &= T(x)\left(\frac{\partial \eta}{\partial x}\right)^2 \end{aligned}\right\} \tag{27}$$

Since these are similar to each other we will take the $+$ sign in one case, and the $-$ sign in the other when taking the square root. Then we have

$$\left.\begin{aligned} \frac{\partial \xi}{\partial t} - \sqrt{\frac{T(x)}{\mu(x)}}\,\frac{\partial \xi}{\partial x} &= 0 \\ \frac{\partial \eta}{\partial t} + \sqrt{\frac{T(x)}{\mu(x)}}\,\frac{\partial \eta}{\partial x} &= 0 \end{aligned}\right\} \tag{28}$$

from which we get the equations of our transformation

$$\left.\begin{aligned} \xi &= t + \int \sqrt{\frac{T}{\mu}}\,dx \\ \eta &= t - \int \sqrt{\frac{T}{\mu}}\,dx \end{aligned}\right\} \tag{29}$$

Substituting in the original equation we now get for the differential equation of the string in these coordinates

$$\frac{\partial^2 y}{\partial \xi \partial \eta} = \frac{1}{8}\psi(\xi, \eta)\left(\frac{\partial y}{\partial \xi} - \frac{\partial y}{\partial \eta}\right), \tag{30}$$

where

$$\psi(\xi, \eta) = \left(\frac{\partial}{\partial \xi} - \frac{\partial}{\partial \eta}\right) \log \{\mu[\phi(\xi, \eta)]T[\phi(\xi, \eta)]\} \tag{31}$$

Here $\phi(\xi, \eta) = x$, and is gotten from

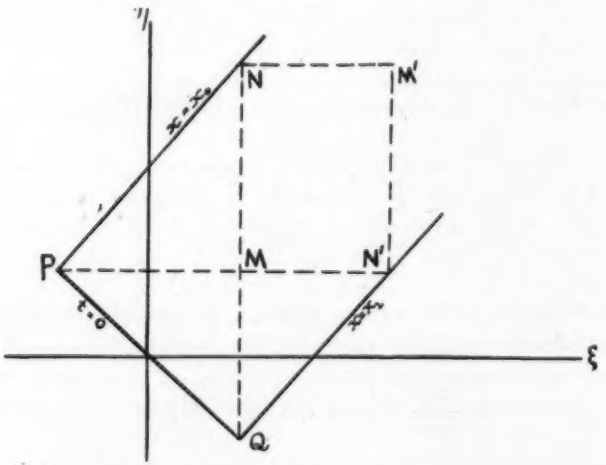
$$\int \sqrt{\frac{T}{\mu}}\, dx = \frac{\xi - \eta}{2}. \tag{32}$$

It might be of interest to state in passing that equation (30) is of the type satisfied by each of the Cartesian coordinates of a point on a surface, expressed in terms of the parameters of conjugate directions on the surface.

The invariants are easily seen to be

$$\left.\begin{aligned} h &= \frac{\partial}{\partial \xi}\psi(\xi, \eta) - \psi^2(\xi, \eta) \\ k &= -\frac{\partial}{\partial \eta}\psi(\xi, \eta) - \psi^2(\xi, \eta) \end{aligned}\right\} \tag{33}$$

It is evident that neither of these will vanish except possibly for a very special function $\psi(\xi, \eta)$, so an immediate general solution is not apparent. We will obtain a series development by the method of Picard.[8]



The figure shows the region of integration in the new coordinates. The x axis has been rotated through -45 degrees about the origin, and the lines $x = x_0$ and $x = x_1$ take the positions shown, all three remaining straight. We divide the whole region into smaller ones as shown by means of parallels to the ξ axis and the η axis. Let us first consider region (1). Along the $t = 0$ line we know the initial position of the string as $y = f(t, x) = f(\xi + \eta, \xi - \eta)$.

Let

$$\frac{\partial^2 y}{\partial \xi \partial \eta} = \lambda \psi \left(\frac{\partial y}{\partial \xi} - \frac{\partial y}{\partial \eta} \right), \tag{34}$$

and

$$y = y_0 + \lambda y_1 + \lambda^2 y_2 + \cdots + \lambda^n y_n + \cdots, \tag{35}$$

Then

$$\frac{\partial^2 y_0}{\partial \xi \partial \eta} + \lambda \frac{\partial^2 y_1}{\partial \xi \partial \eta} + \cdots = \psi \left[\lambda \left(\frac{\partial y_0}{\partial \xi} - \frac{\partial y_0}{\partial \eta} \right) + \lambda^2 \left(\frac{\partial y_1}{\partial \xi} - \frac{\partial y_1}{\partial \eta} \right) + \cdots \right]. \tag{36}$$

Therefore

$$\frac{\partial^2 y_n}{\partial \xi \partial \eta} = \psi \left(\frac{\partial y_{n-1}}{\partial \xi} - \frac{\partial y_{n-1}}{\partial \eta} \right). \tag{37}$$

We now let $\partial^2 y_0 / \partial \xi \partial \eta = 0$, whence $y_0 = f_1(\xi + \eta) + f_2(\xi - \eta)$, the initial position of the string.

It is evident that y_0 reduces to $\varphi_1(\eta)$ for $\xi = \xi_0$, to $\varphi_2(\xi)$ for $\eta = \eta_0$, and that $\varphi_1(\xi_0)$ must equal $\varphi_2(\eta_0)$. The other y_1, y_2, and so on can then be placed equal to 0 for $\xi = \xi_0$, ξ_1, and for $\eta = \eta_0$, η_1.

Then from (37) we have the formula for the y's

$$y_n(\xi, \eta) = \int_{\xi_0}^{\xi} \partial\mu \int_{\eta_0}^{\eta} \left[\psi(\mu, \nu) \frac{\partial y_{n-1}}{d\mu} - \psi(\mu, \nu) \frac{\partial y_{n-1}}{\partial \nu} \right] d\nu. \tag{38}$$

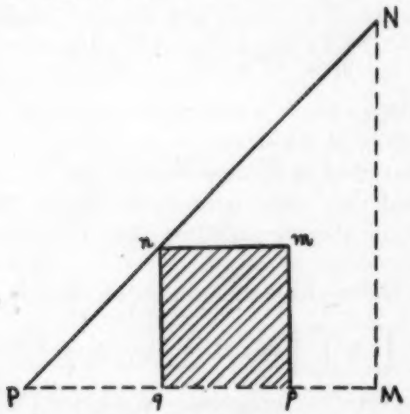
The convergence and uniqueness of solution (35) have been proved by Picard, as above, and will not be repeated here. Even though $\varphi_1(\eta)$ and $\varphi_2(\xi)$ or their derivatives have a finite number of discontinuities, the y's are regular, and $y(\xi, \eta)$ as given by (35) is regular except at a finite number of points.

The boundary conditions may be cared for by the method of Goursat [9] as follows. Take the integral

$$y_n(\xi, \eta) = \int\int_{(PMQ)} \psi(\mu, \nu) \left[\frac{\partial y_{n-1}}{\partial \mu} - \frac{\partial y_{n-1}}{\partial \nu} \right] d\mu d\nu. \quad (39)$$

As shown in the reference above this is solvable with our given initial position of the string, for the requisite number of derivatives will exist, as will be shown in the next chapter. Hence the solution of the equation exists for region (1).

In either of the regions marked (2) the problem is the same. In the one we know y and its derivatives along PM, in the other along MQ, from the solution of region (1). In the first we also know y and all time derivatives, since they are zero, along PM, and in the second along QN'. Consider the upper of the two regions. We have a mixed problem (see Goursat above, p. 119 ff). Again we use the series (35) and the formula of recurrence (38). We set here $y_0 = f_3(\xi + \eta) + f_4(\xi - \eta)$, the value of the integral along PM, where $\eta = \text{const.}$



The boundary conditions come in through the region of integration, which is now $mnqp$. This gives us

$$y_n(\xi, \eta) = \int\int_{(mnqp)} \psi(u, \nu) \left[\frac{\partial y_{n-1}}{\partial \mu} - \frac{\partial y_{n-1}}{\partial \nu} \right] d\mu d\nu. \quad (40)$$

Using these y's the series (35) for this region is convergent as before, as will be the series for the partial derivatives. These series will satisfy the Cauchy conditions, and insure the solution of our equation in regions (2), (Goursat, ib. p. 155 ff).

As for region (3) we have the values of the integral along MN and MN', so the value everywhere is obtained in the same general manner as for (1) and (2) by integrating over the region $MNM'N'$.

Thus the solution exists everywhere so long as the initial position of the string as given by $y = f(\xi + \eta,\ \xi - \eta)$ is sufficiently differentiable. As before stated, this condition will be shown to be met in the following chapter.

Having shown the existence of the solution, we know from the results of Chapter II that it is almost periodic in the mean, and from Chapter I we see that it may be approximated by means of a polynomial of the form

$$P_1(\xi + \eta,\ \xi - \eta) = \sum_1^N A_k(\xi - \eta)e^{i\lambda_k(\xi+\eta)}.$$

Or, referring back to the original axes, the solution may be approximated by means of the polynomial

$$P_2(t,\ x) = \sum_1^N B_k(x)e^{i\lambda_k t}.$$

CHAPTER IV. THE COMPLETION OF THE PROBLEM, AND THE CLOSURE OF THE EIGENFUNKTIONEN OR CHARACTERISTIC FUNCTIONS

We have been working with the transformed equation, but the transformation being a simple one we can return at once to the original form, and will work with it from this point.

From equation (1) we get the corresponding integral equation

$$\int_{t_0}^{t}\int_{t_0}^{t}\left[y(t,\ x) - y(t,\ x_0) - \left(\frac{\partial y}{\partial x}\right)_{x_0}(x - x_0)\right]dt^2$$
$$+ \int_{t_0}^{t}\int_{t_0}^{t}\int_{x_0}^{x} y\,\frac{\partial \log T}{\partial x}\,dx dt^2 - \int_{t_0}^{t}\int_{t_0}^{t}\int_{x_0}^{x}\int_{x_0}^{x} y\,\frac{\partial^2 \log T}{\partial x^2}\,dx^2 dt^2$$
$$= \int_{x_0}^{x}\int_{x_0}^{x}\frac{\mu}{T}\left[y(t,\ x) - y(t_0,\ x) - \left(\frac{\partial y}{\partial t}\right)_{t_0}(t - t_0)\right]dx^2. \qquad (41)$$

Now any solution whatever of the differential equation (1) is also a solution of this, its associated integral equation. Therefore the approximating trigonometric polynomial solution of the differential equation, arising from the almost periodic in the mean character of y, is a solution of (41). In addition we know from the theory of integral equations that the uniform limit of a number of solutions of (41) is also a solution of it. We now wish to show that each separate term of the approximating polynomial solution is also a solution.

Since $f(t, x)$ is a solution of (41), it is evident that $f(t + \tau, x)$ is also one, where τ is a displacement vector, and since a constant times a solution is also a solution, we know that $k(\tau)f(t + \tau, x)$ will satisfy (41).

Now from the uniformity properties of functions almost periodic in the mean, and from the existence of the mean value we also have the following solutions of (41)

$$\left.\begin{aligned} &\int_{-A}^{A} k(\tau)f(t + \tau, x)dt \\ &\text{l.i.m.}_{A\to\infty} \frac{1}{2A}\int_{-A}^{A} k(\tau)f(t + \tau, x)dt \end{aligned}\right\} \tag{42}$$

Using these solutions, and referring to pages 180–181 of Chapter I, we see that the functions $f_a(t, x)$ and $f^{(a)}(t, x)$ will also satisfy the integral equation. Then since the uniform limit of solutions is a solution

$$M_{\tau}[f^{(a)}(\tau, x)e^{i\Lambda(t-\tau)}] \tag{43}$$

satisfies (41).

Considering this expression we see that it is really but a single term, for the averaging process causes everything to vanish except a single term corresponding to the Λ, the $f^{(a)}(t, x)$ being considered as a trigonometric series as on page 181.

Since $f(t, x)$ and $f^{(a)}(t, x)$ have been shown to differ in the mean by a vanishingly small quantity, we may take a single term of the polynomial approximation of $f(t, x)$ for our purpose. Therefore $A_k(x)e^{i\Lambda_k t}$ satisfies (41).

We now wish to show that these single terms of the polynomial also satisfy our differential equation (1).

Putting $A_k(x)e^{i\lambda_k t}$ in (41) we see that $e^{i\lambda_k t}$ gives no trouble, since it can be differentiated as often as we please when we reduce (41) back to (1) by differentiating twice with respect to each t and x. The only question is as to the differentiability of $y(t, x)$ with respect to x. This is the same question we have met before concerning the initial position, for it is simply a special case for $t = 0$ of this, the general case. The question of the two-fold differentiability of $\log T(x)$ with respect to x arises in the integral equation, and we must assume this to be possible. But we can prove the existence of the required derivatives of $y = f(t, x)$ as follows.

Differentiate (41) twice with respect to t. This gives the equation

$$y(t, x) = -y(t, x_0) + \left(\frac{\partial y}{\partial x}\right)_{x_0}(x - x_0)$$

$$- \int_{x_0}^{x} y \frac{\partial \log T}{\partial x}\, dx + \int_{x_0}^{x}\int_{x_0}^{x} y \frac{\partial^2 \log T}{dx^2}\, \partial x^2$$

$$+ \int_{x_0}^{x}\int_{x_0}^{x} \frac{\mu}{T}\left[y(t, x) - y(t_0, x) - \left(\frac{\partial y}{\partial t}\right)_{t_0}(t - t_0)\right] dx^2 \qquad (44)$$

where everything but $y(t, x)$ has been placed on the right. It is now perfectly evident that the right side can be differentiated once with respect to x. Thus $\partial y/\partial x$ exists.

Now the first term on the right is a function of t alone, the second will cause no trouble, and the two double integrals are immediately twice differentiable with respect to x. Hence it is only the single integral that we must examine now.

Since we have shown the existence of $\partial y/\partial x$, we can write

$$y = \int_{x_0}^{x} \frac{\partial y}{\partial x} dx + \left(\frac{\partial y}{\partial x}\right)_{x_0} \qquad (45)$$

and substituting this value, the single integral becomes

$$\int_{x_0}^{x}\left[\int_{x_0}^{x} \frac{\partial y}{\partial x} dx + \left(\frac{\partial y}{\partial x}\right)_{x_0}\right] \frac{\partial \log T}{\partial x} dx \qquad (46)$$

which is twice differentiable with respect to x. The process could be continued, but we need only the two derivatives, so we see that the initial position and the general integral of our system must possess the requisite derivatives.

Thus single terms of the polynomial $P(t, x)$ are solutions of the differential equation.

As a result we see that we can approximate uniformly in the mean to a solution of (1), reducing to $\phi(x)$ for $t = 0$, by means of $\sum_1^N H_k(x)e^{i\Lambda_k t}$, where $H_k(x)e^{i\Lambda_k t}$ satisfies (1).

Now the functions $H_k(x)$ are the eigenfunktionen or the characteristic functions of the system, and satisfy the differential equation of eigenfunktionen

$$\frac{\partial}{\partial x}\left[T(x)\frac{\partial y}{\partial x}\right] = -\mu\Lambda_k^2 y \tag{47}$$

Then we see that the initial position of the string can be approximated in the mean as closely as desired by a denumerable set of eigenfunktionen, since we may take

$$\int_{x_0}^{x_1}\left[\phi(x) - \sum_1^N H_k(x)\right]^2 dx \tag{48}$$

so as to be less than ϵ, $\epsilon/2$, $\epsilon/4$, etc., by properly choosing N each time.

Now take any closed, denumerable set $\phi_k(x)$ of initial positions. Each ϕ can be approximated by a denumerable set of eigenfunktionen as above. Since any quadratically summable function can be approximated by a series of sines and cosines, we can consider our $\phi_k(x)$ thus expressed. Then taking the whole set ϕ_k we have a denumerable set of denumerable sets of eigenfunktionen. From this we immediately conclude that any quadratically initial position of the string is expressible by means of a denumerable set of eigenfunktionen of the system. For the present problem, the eigenfunktionen are the standing waves of the string.

BIBLIOGRAPHY

1. H. Bohr. "Zur Theorie der fast periodischen Funktionen." *Acta. Math.*, 45: pp. 29–127 (1924).
2. H. Bohr. "Zur Theorie der fast periodischen Funktionen." *Acta. Math.*, 46–47 (1925–1926).
3. S. Bochner. "Beitrage zur Theorie der fast periodischen Funktionen." *Math. Ann.*, 96: pp. 119–147 (1926).
4. P. Franklin. "The Elementary Theory of Almost Periodic Functions of two Variables." Jour. Math. and Phys., M. I. T., Vol. V, No. 1, pp. 40–54 (1926).
   "The Fundamental Theorem of Almost Periodic Functions of two Variables." *Ib.*, No. 4, pp. 201–237.
5. H. Weyl. "Beweis des Fundamentalsatzes in der Theorie der fast-periodischen Funktionen (aus einem Schreiben an Herrn Harald Bohr)." *Sitzber. der preuss. Akad. der Wiss.* (Phys.-Math. Klasse), 18, 19: pp. 211–214 (1926).
   "Integralgleichungen und fastperiodischen Funktionen." *Math. Ann.*, 97: pp. 338–356 (1926).
6. W. Stepanoff. "Über einige Verallgemeinerungen der fast periodischen Funktionen." *Math. Ann.*, 95: pp. 473–498 (1926).
7. N. Wiener. "The Spectrum of an Arbitrary Function." *Proc. Lond. Math. Soc.*, Ser. 2, Vol. 27, Part 7, pp. 483–496 (1926).
8. M. Picard. "Mémoire sur la théorie des équations aux dérivées partielles." *Jour. de Math. Pure et Appl.*, Ser. 4, No. 6, pp. 19–20 (1890).
9. E. Goursat. "Cours d'Analyse Mathématique." Ed. 2, Tome III, p. 114 ff.
10. S. Bochner. "Sur les fonctions presque-périodiques de Bohr." *Comptes Rendus*, t. 180: p. 1156 (1925).
11. H. Bohn and O. Neugebauer. "Über lineare Differentialgleichungen mit konstanten Koeffizienten und fastperiodischen rechter Seite." *Sitzber. Gott.*, 16 July, 1926, pp. 1–13.
12. E. Eschangon. "Nouvelles recherches sur les fonctions quasipériodiques." *Ann. de l'Observ. de Bordeaux*, 1919.
13. J. Favard. "Sur les fonctions harmoniques presque-périodiques." These, Paris, 1927.
14. J. Favard. "Sur les équations différentielles à coefficients presque-périodiques." *Acta. Math.*, 51: p. 31 (1927).
15. J. Favard. "Sur les fonctions harmoniques presque-périodiques." *Comptes Rendus*, 182: p. 757 (1926).
16. S. Bochner. "Properties of Fourier series of almost periodic functions." *Proc. Lond. Math. Soc.*, 26, Ser. 2, p. 433 (1925).
17. H. Bohr. "Sur une classe de transcendantes entières." *Comptes Rendus*, 181: p. 766 (1925).
18. H. Bohr. "Über Fourrierreihen fastperiodischen Funktionen." *Math. Zeitschr.*, 23: pp. 38–44 (1925).

AN APPLICATION OF THE INVARIANT AREA PROPERTIES OF ALGEBRAIC POLYNOMIALS TO THE DERIVATION OF FORMULAS FOR MECHANICAL INTEGRATION *

By Prescott D. Crout

1. Introduction. The purpose of this paper is to investigate the invariant area properties of algebraic polynomials, and to apply these properties to the derivation of formulas for approximate integration. It will be shown that there exists a family of polynomials analogous to and including the polynomials of Legendre, such that the roots of each fix the spacing of the points of an integration formula which satisfies certain conditions, and which gives exact results when applied to any polynomial of degree smaller than a given number. The formulas of Gauss follow directly as a special application of the general theory.

In this paper only algebraic polynomials will be considered, no attempt being made to investigate the error involved when a formula of mechanical integration is used to approximate the integral of an arbitrary function.

2. Formulas for the Integral of a Polynomial Over a Given Interval. Let there be given the n points (c_1, y_1), (c_2, y_2), (c_3, y_3), $\cdots$, (c_n, y_n); then the polynomial of lowest degree passing through these points is

$$P(x) = \sum_{i=1}^{n} \frac{\prod_{\substack{j=1 \\ j \neq i}}^{n} (x - c_j)}{\prod_{\substack{j=1 \\ j \neq i}}^{n} (c_i - c_j)} y_i, \tag{1}$$

* This paper was prepared under the supervision of Professor George Rutledge, of the Massachusetts Institute of Technology, who offered many valuable suggestions.

which is the formula of Lagrange.[1] We shall now integrate $P(x)$ over a given interval, which can without loss of generality be taken as the interval $(-\theta, \theta)$, since this follows from a particular choice of the origin in (1). We thus have

$$\int_{-\theta}^{\theta} P(x)dx = \int_{-\theta}^{\theta} \sum_{i=1}^{n} \frac{\prod_{\substack{j=1 \\ j\neq i}}^{n} (x - c_j)}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (c_i - c_j)} y_i dx,$$

or

$$\int_{-\theta}^{\theta} P(x)dx = \sum_{i=1}^{n} y_i \frac{\int_{-\theta}^{\theta} \prod_{\substack{j=1 \\ j\neq i}}^{n} (x - c_j)dx}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (c_i - c_j)}; \tag{2}$$

hence

$$\int_{-\theta}^{\theta} P(x)dx = \sum_{i=1}^{n} K_i y_i \tag{3}$$

where

$$K_i = \frac{\int_{-\theta}^{\theta} \prod_{\substack{j=1 \\ j\neq i}}^{n} (x - c_j)dx}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (c_i - c_j)}. \tag{4}$$

Here it is to be noted that the K's can be determined once and for all when θ and the c's are given; thus the K's are independent of $P(x)$. By applying (3) to the polynomial P = constant, we find that

$$\sum_{i=1}^{n} K_i = 2\theta; \tag{5}$$

also, from (4) it is evident that when the points are symmetrically

[1] *Crout: A Method for Deriving Formulas of Interpolation*, this *Journal*, vol. 8, p. 29.

placed about the origin, or $-c_j = c_{n-j+1}$, then

$$K_j = K_{n-j+1} \quad j = 1, 2, 3, \cdots, \frac{n}{2}, \; n \text{ even}, \qquad j = 1, 2, 3, \cdots, \frac{n-1}{2}, \; n \text{ odd}. \tag{6}$$

For convenience we shall now leave θ arbitrary and fix the c's by their ratios to θ; thus

$$c_i = r_i\theta \quad i = 1, 2, 3, \cdots, n. \tag{7}$$

Substituting (7) in (4) and making the substitution

$$u\theta = x,$$

(4) becomes

$$K_i = \frac{\theta \int_{-1}^{1} \prod_{\substack{j=1 \\ j\neq i}}^{n} (u - r_j)du}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (r_i - r_j)};$$

thus (3) may be written

$$\int_{-\theta}^{\theta} P(x)dx = \theta \sum_{i=1}^{n} k_i y_i \tag{8}$$

where

$$k_i = \frac{\int_{-1}^{1} \prod_{\substack{j=1 \\ j\neq i}}^{n} (u - r_j)du}{\prod_{\substack{j=1 \\ j\neq i}}^{n} (r_i - r_j)}. \tag{9}$$

Obviously

$$K_i = \theta k_i \quad i = 1, 2, 3, \cdots, n; \tag{10}$$

thus from (5) and (10) we see that

$$\sum_{i=1}^{n} k_i = 2; \tag{11}$$

and from (6) and (10) we see that when the points are symmetrically placed about the origin, or $-r_j = r_{n-j+1}$, then

$$k_j = k_{n-j+1} \quad j = 1, 2, 3, \cdots, \frac{n}{2}, \ n \text{ even}, \qquad j = 1, 2, 3, \cdots, \frac{n-1}{2}, \ n \text{ odd}. \tag{12}$$

As an example let us suppose that the n points are placed so as to to divide the interval $(-\theta, \theta)$ into $n-1$ equal subintervals; then the ratios are

$$r_i = \frac{c_i}{\theta} = \frac{-\theta + \dfrac{2\theta(i-1)}{n-1}}{\theta} = \frac{2i-n-1}{n-1} \quad i = 1, 2, 3, \cdots, n,$$

and

$$k_i = \frac{\displaystyle\int_{-1}^{1} \prod_{\substack{j=1 \\ j \neq i}}^{n} ([n-1]u + n + 1 - 2j)du}{2^{n-1} \displaystyle\prod_{\substack{j=1 \\ j \neq i}}^{n} (i-j)}. \tag{13}$$

In this case the formulas (8) are those of Cotes, and the constants (13) are Cotes' coefficients. For example, if $n = 5$, (8) becomes

$$\int_{-\theta}^{\theta} P(x)dx = \frac{\theta}{45}\{7y_1 + 32y_2 + 12y_3 + 32y_4 + 7y_5\}.$$

3. Invariant Area. Let us now suppose that the distribution of the points is such that one of the K's, K_i to fix ideas, is zero. This is obviously the same as saying that the r's are distributed in such a way that k_i vanishes. Omitting the ith point, it is seen that an infinite number of polynomials of degree $n-1$ can be passed through the remaining $n-1$ points; however, in the case of all of these polynomials $\displaystyle\int_{-\theta}^{\theta} P(x)dx$ remains

invariant, for y_i can be taken as the parameter of the family, and the term $\theta k_i y_i$ contributes nothing to the sum (8) since k_i is zero. Since $P(x)$ can, in particular, be taken as the polynomial of degree $n-2$ that is determined by the remaining $n-1$ points, it is evident that the k's are fixed independent of r_i; thus in this case

$$\int_{-\theta}^{\theta} P(x)dx = \theta \sum_{\substack{j=1 \\ j \neq i}}^{n} k_j y_j. \tag{14}$$

If, now, a new formula of integration is derived which is based upon these same $n-1$ points and upon an arbitrary point c_m of ratio r_m, we have from (8)

$$\int_{-\theta}^{\theta} P(x)dx = \theta k_m y_m + \theta \sum_{\substack{j=1 \\ j \neq m}}^{n} k_j y_j; \tag{15}$$

however, since the k's are uniquely determinate, and since (15) is satisfied identically if $k_m = 0$ and the other k's of (15) are identical with the corresponding k's of (14), it follows that (14) and (15) must be identical; and $k_m = 0$ regardless of r_m.

This relation also follows directly from (9), for the condition that k_i vanish is

$$k_i = \frac{\displaystyle\int_{-1}^{1} \prod_{\substack{j=1 \\ j \neq i}}^{n} (u - r_j)du}{\displaystyle\prod_{\substack{j=1 \\ j \neq i}}^{n} (r_i - r_j)} = 0, \tag{16}$$

or

$$\int_{-1}^{1} \prod_{\substack{j=1 \\ j \neq i}}^{n} (u - r_j)du = 0, \tag{17}$$

since the denominator of (16) is finite. The condition (17) does not involve r_i; hence if k_i vanishes for a particular value of r_i, it vanishes for all values of r_i, the other points being held fixed.

Let us now consider k_h where $h = 1, 2, 3, \cdots, n$, but where $h \neq i$. We have from (9)

$$\frac{\partial k_h}{\partial r_i} = \frac{\partial}{\partial r_i}\left\{\frac{\int_{-1}\prod_{\substack{j=1\\ j\neq h}}^{n}(u - r_j)du}{\prod_{\substack{j=1\\ j\neq h}}^{n}(r_h - r_j)}\right\}$$

$$= \frac{\left\{\prod_{\substack{j=1\\ j\neq h}}^{n}(r_h - r_j)\right\}\left\{\int_{-1} - \prod_{\substack{j=1\\ j\neq i,h}}^{n}(u - r_j)du\right\} - \left\{-\prod_{\substack{j=1\\ j\neq i,h}}^{n}(r_h - r_j)\right\}\left\{\int_{-1}\prod_{\substack{j=1\\ j\neq h}}^{n}(u - r_j)du\right\}}{\left\{\prod_{\substack{j=1\\ j\neq h}}^{n}(r_h - r_j)\right\}^2}$$

$$= \frac{\int_{-1}\prod_{\substack{j=1\\ j\neq i}}^{n}(u - r_j)du}{(r_h - r_i)\prod_{\substack{j=1\\ j\neq h}}^{n}(r_h - r_j)};$$

thus finally it follows that

$$\frac{\partial k_h}{\partial r_i} = \left\{-\frac{\prod_{\substack{j=1\\ j\neq i,h}}^{n}(r_i - r_j)}{\prod_{\substack{j=1\\ j\neq h}}^{n}(r_h - r_j)}\right\}k_i. \tag{18}$$

Since the coefficient of k_i in (18) is finite but not zero, it follows that the necessary and sufficient condition for k_h to be independent of r_i is that k_i vanish, which condition is (17) and hence independent of r_i.

Replacing n and i of (17) by $n + 1$, noting (14), we see that *if $P(x)$ is any polynomial of degree n passing through n fixed*

points, $\int_{-a}^{a} P(x)dx$ *is invariant and given by* (8) *if and only if*

$$\int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du = 0. \tag{19}$$

As an illustration, let us consider n points placed symmetrically about the origin; then if n is even, we have

$$\int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du = \int_{-1}^{1} \prod_{j=1}^{n/2} (u^2 - r_j^2)du$$

which may or may not vanish; however, if n is odd, we have

$$\int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du = \int_{-1}^{1} u \prod_{j=1}^{(n-1)/2} (u^2 - r_j^2)du = 0$$

since $u \prod_{j=1}^{(n-1)/2} (u^2 - r_j^2)$ is an odd function; hence in this case, where n is odd, (19) is satisfied and $\int_{-a}^{a} P(x)dx$ is invariant and given by (8) for any polynomial of degree n passing through these n points. As an application we may note that Cotes' formula for n points gives exact results when applied to any polynomial of degree n if n is odd.

In the preceding illustration we merely verified the fact that (19) is satisfied in the case of an odd number of points symmetrically placed about the origin; however, it is obvious that (19) may be taken as a condition subject to which the n points are initially chosen, the resulting formulas all giving exact results when applied to any polynomial of degree n. For example, if $n = 2$ we have

$$\int_{-1}^{1} (u - r_1)(u - r_2)du = 0,$$

or

$$r_1 r_2 = -\frac{1}{3}; \tag{20}$$

also

$$k_1 = \frac{\int_{-1}^{1}(u - r_2)du}{r_1 - r_2} = \frac{2r_2}{r_2 - r_1} = \frac{2}{1 - r_1/r_2},$$
$$k_2 = \frac{\int_{-1}^{1}(u - r_1)du}{r_2 - r_1} = \frac{-2r_1}{r_2 - r_1} = \frac{2}{1 - r_2/r_1}. \tag{21}$$

Taking r_1 as a parameter, we have

$$r_2 = -\frac{1}{3r_1}, \quad k_1 = \frac{2}{1 + 3r_1^2}, \quad k_2 = \frac{2}{1 + 1/3r_1^2}.$$

We may also verify (14) and derive (21) by considering three points and placing $k_3 = 0$, in which case

$$r_1 r_2 = -\frac{1}{3}$$

as in (20). Then

$$k_1 = \frac{\int_{-1}^{1}(u - r_2)(u - r_3)du.}{(r_1 - r_2)(r_1 - r_3)} = \frac{2(1 + 3r_2r_3)}{3r_1^2 + 1 - 3r_1r_3 + 3r_2r_3}$$
$$k_1 = \frac{2}{1 + \dfrac{r_1(3r_1 - 3r_3)}{1 + 3r_2r_3}} = \frac{2}{1 - \dfrac{r_1}{r_2}}; \tag{22}$$

similarly

$$k_2 = \frac{2}{1 - \dfrac{r_2}{r_1}};$$

thus the results (22) are independent of r_3, as is required by (18), and are identical with (21). If it is desired that the points be symmetrical about the origin, we have

$$r_1 = -r_2,$$

which with (20) gives

$$-r_1 = r_2 = \frac{1}{\sqrt{3}};$$

also, from (21)

$$k_1 = k_2 = 1,$$

and finally

$$\int_{-\theta}^{\theta} P(x)dx = \theta P\left(-\frac{\theta}{\sqrt{3}}\right) + \theta P\left(\frac{\theta}{\sqrt{3}}\right) \tag{23}$$

for any polynomial of second degree. Since

$$\int_{-\theta}^{\theta} x^{2i+1}dx = 0 \quad i = 0, 1, 2, 3, \cdots,$$

(23) gives exact results when applied to any polynomial of third degree.

4. Invariancy of Higher Order. We shall now investigate the conditions under which the integration formula (8) gives exact results when applied to any polynomial of degree $n + i$, in which case we define i as the *order of invariancy* of the formula. We thus see that the invariancy considered in Section 3 was that of order zero; obviously $i \geq 0$ unless there is no invariancy, in which case $i = -1$.

It has been shown in Section 3, (15) and (19), that the integration formula for the n points $r_1, r_2, \cdots, r_n$ is identical with the integration formula for these points and some arbitrary point r_{n+1} if and only if

$$\int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du = 0. \tag{24}$$

Similarly, the integration formula for the $n + 1$ points $r_1, r_2, \cdots, r_{n+1}$ is identical with that for these points and another arbitrary point r_{n+2} if and only if

$$\int_{-1}^{1} \prod_{j=1}^{n+1} (u - r_j)du = 0,$$

or

$$\int_{-1}^{1} u \prod_{j=1}^{n} (u - r_j)du - r_{n+1} \int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du = 0; \tag{25}$$

hence the formula for the n points $r_1, r_2, \cdots, r_n$ is identical with the formula for the $n+2$ points $r_1, r_2, \cdots, r_{n+2}$ if and only if (24) and (25) are both satisfied, or if

$$\int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du = 0,$$

$$\int_{-1}^{1} u \prod_{j=1}^{n} (u - r_j)du = 0.$$

Proceeding as with (25) we see that the formula for the n points $r_1, r_2, \cdots, r_n$ is identical with that for the $n+i+1$ points $r_1, r_2, \cdots, r_{n+i+1}$, the points $r_{n+1}, r_{n+2}, \cdots, r_{n+i+1}$ being arbitrary, if and only if [2]

$$\int_{-1}^{1} u^\alpha \prod_{j=1}^{n} (u - r_j)du = 0, \quad \alpha = 0, 1, 2, \cdots, i. \tag{26}$$

But the integration formula for $n+i+1$ points gives exact results when applied to any polynomial of degree $n+i$; hence from (26) it is evident that the integration formula (8) for the n points $r_1, r_2, \cdots, r_n$ has invariancy of order i if and only if [3]

[2] More explicitly, the formula for the n points $r_1, r_2, \cdots, r_n$ is identical with that for these points and $i+1$ arbitrary points $r_{n+1}, r_{n+2}, \cdots, r_{n+i+1}$ if and only if the $i+1$ conditions

$$\int_{-1}^{1} \prod_{j=1}^{n+\alpha} (u - r_j)du = 0, \quad \alpha = 0, 1, 2, \cdots i \tag{I}$$

are satisfied. But (I) may be written

$$a_{\alpha,0} \int_{-1}^{1} \prod_{j=1}^{n} (u - r_j)du + a_{\alpha,1} \int_{-1}^{1} u \prod_{j=1}^{n} (u - r_j)du + a_{\alpha,2} \int_{-1}^{1} u^2 \times \prod_{j=1}^{n} (u - r_j)du + \cdots + a_{\alpha,\alpha} \int_{-1}^{1} u^\alpha \prod_{j=1}^{n} (u - r_j)du = 0, \quad \alpha = 0, 1, 2, \cdots i, \tag{II}$$

where the a's involve $r_{n+1}, r_{n+2}, \cdots r_{n+i+1}$; also it is evident that in each condition (II), all terms but the last vanish due to the preceding conditions; thus we have (26).

[3] The conditions (27) also follow directly from (17) and (18), for the formula based upon the $n+i+1$ points $r_1, r_2, \cdots, r_{n+i+1}$ is independent of $r_{n+1}, r_{n+2}, \cdots, r_{n+i+1}$ if and only if the $i+1$ conditions

$$\int_{-1}^{1} \prod_{\substack{j=1 \\ j \neq n+m}}^{n+i+1} (u - r_j)du = 0, \quad m = 1, 2, 3, \cdots, i+1; \tag{III}$$

$$\int_{-1}^{1} Q_i(u) \prod_{j=1}^{n} (u - r_j) du = 0 \tag{27}$$

where $Q_i(u)$ represents any arbitrary polynomial of degree i. Furthermore since there cannot in general be more equations of condition than there are undetermined constants, or r's, we have in general

$$i \leq n - 1. \tag{28}$$

For example let us again consider the case where $n = 2$; then from (28) we have

$$i \leq 1, \quad \text{or} \quad i = 0, 1.$$

If $i = 0$; then from (20) it is only necessary that

$$r_1 r_2 = -\frac{1}{3};$$

however, if $i = 1$, then it is also necessary that

$$\int_{-1}^{1} u(u - r_1)(u - r_2) du = 0,$$

or

$$r_1 + r_2 = 0,$$

due to (27). Thus if $i = 1$, then $-r_1 = r_2 = 1/\sqrt{3}$; and we see that the formula (23) has invariancy of order one, which again proves that (23) gives exact results when applied to any polynomial of third degree.

5. Invariancy of Order i. Formulas of Gauss. We have seen that the $i + 1$ conditions (26) for invariancy of order i can be represented by the single expression (27); however, it is evident that this is a simplification of notation alone. We shall now derive a family of polynomials of degree n such that the but (III) may be written

$$\int_{-1}^{1} G_{m,i}(u) \prod_{j=1}^{n} (u - r_j) du = 0, \quad m = 1, 2, 3, \cdots, i + 1, \tag{IV}$$

where $G_{m,i}(u)$ is a polynomial in u of degree i; thus it is obvious that (27) is a sufficient condition that (IV) be satisfied. Since the coefficients in the polynomials $G_{m,i}(u)$ involve only the arbitrary quantities $r_{n+1}, r_{n+2}, \cdots, r_{n+i+1}$, it is evident that (27) is also a necessary condition.

n roots of any one of them can be taken as a set of r's which satisfies (27), and such that any n r's which satisfy (27) are the roots of one of these polynomials.

Let us place

$$\prod_{j=1}^{n}(u - r_j) = \frac{d^{i+1}}{du^{i+1}}R(u) \tag{29}$$

where $R(u)$ is a polynomial in u of degree $n + i + 1$. It is evident that (29) does not fix $R(u)$, and that the constants of integration can be determined so that $R(u)$ satisfies the conditions

$$R(-1) = 0, \quad R^1(-1) = 0, \quad R^2(-1) = 0, \quad \cdots, \quad R^i(-1) = 0 \tag{30}$$

in addition to (29); thus $R(u)$ is completely determined. Substituting (29) in (27) and integrating by parts successively, we have

$$\int_{-1}^{1} Q_i(u)\prod_{j=1}^{n}(u - r_j)du = [Q_i(u)R^i(u) - Q_i^1(u)R^{i-1}(u) + \cdots \pm Q_i^i(u)R(u)]_{-1}^{1} = 0,$$

since $Q_i^{i+1}(u) = 0$. Hence from (30) we have

$$\int_{-1}^{1} Q_i(u)\prod_{j=1}^{n}(u - r_j)du = Q_i(1)R^i(1) - Q_i^1(1)R^{i-1}(1) + \cdots \pm Q_i^i(1)R(1) = 0;$$

and since $Q_i(u)$ is arbitrary, it is necessary that

$$R(1) = 0, \quad R^1(1) = 0, \quad R^2(1) = 0, \quad \cdots, \quad R^i(1) = 0; \tag{31}$$

thus from (30) and (31) we see that

$$R(u) = (u^2 - 1)^{i+1}\prod_{j=1}^{n-i-1}(u - \beta_j) \tag{32}$$

where $\beta_1, \beta_2, \cdots, \beta_{n-i-1}$ represent the other roots of $R(u)$. From (32) and (29) it follows that

$$\prod_{j=1}^{n}(u - r_j) = \frac{d^{i+1}}{du^{i+1}}\left\{(u^2 - 1)^{i+1}\prod_{j=1}^{n-i-1}(u - \beta_j)\right\}; \tag{33}$$

thus *we have a family of polynomials*

$$\prod = \frac{d^{i+1}}{du^{i+1}} \left\{ (u^2 - 1)^{i+1} \prod_{j=1}^{n-i-1} (u - \beta_j) \right\} \tag{34}$$

of parameters $\beta_1, \beta_2, \cdots, \beta_{n-i-1}$, *such that the n roots of any of them may be taken as a set of n r's which satisfies* (27). Conversely, if a set of n r's is given which satisfies (27); then $\prod_{j=1}^{n} (u - r_j)$ is a polynomial of the family (34), since (33) is a condition which is both necessary and sufficient. In other words *the necessary and sufficient condition for an integration formula* (8) *to have invariancy of order i is that its r's are the roots of one of the polynomials of the family* (34).

As an example we may again consider the case where $n = 2$, $i = 0$; then from (34) we have

$$\prod = \frac{d}{du} \{ (u^2 - 1)(u - \beta_1) \} = 3u^2 - 2u\beta_1 - 1,$$

the roots of which are

$$r_1 = \frac{1}{3}(\beta_1 - \sqrt{\beta_1^2 + 3}), \quad r_2 = \frac{1}{3}(\beta_1 + \sqrt{\beta_1^2 + 3}).$$

Also,

$$r_1 r_2 = \frac{1}{9}(\beta_1^2 - \beta_1^2 - 3) = -\frac{1}{3},$$

which agrees with (20).

We shall now consider the case where i has its greatest value, which from (28) is given by the relation

$$i = n - 1;$$

then (34) becomes

$$\prod = \frac{d^n}{du^n}(u^2 - 1)^n, \tag{35}$$

which except for a constant factor is the Legendre polynomial of degree n. It is thus evident that *a formula of integration for*

n points has its greatest invariancy, $i = n - 1$, when $r_1, r_2, \cdots, r_n$ are the roots of the Legendre polynomial of degree n. Such formulas are known as Gauss' formulas;[4,5] thus Gauss' formulas are those of greatest invariancy.

Thus far the fact that $i = n - 1$ is actually the *greatest* value which i can take, which was stated in (28), was made to depend rather loosely upon (27); however, we are now in a position to see that this is rigorously true, for in order to obtain an invariancy of order $n - 1$, it is necessary to fix n distinct points so that $r_1, r_2, \cdots, r_n$ are the roots of the nth degree polynomial of Legendre.

As an example of Gauss' formulas we may consider the case where $n = 3$; then from (35) we have

$$\prod = \frac{d^3}{du^3}(u^2 - 1)^3 = 24(5u^3 - 3u),$$

the roots of which are

$$r_1 = -\sqrt{\frac{3}{5}}, \quad r_2 = 0, \quad r_3 = \sqrt{\frac{3}{5}};$$

also

$$k_1 = \frac{\int_{-1}^{1} u\left(u - \sqrt{\frac{3}{5}}\right)du}{\left(-\sqrt{\frac{3}{5}}\right)\left(-\sqrt{\frac{3}{5}} - \sqrt{\frac{3}{5}}\right)} = \frac{5}{9} = k_3$$

due to (12). From (11) we have

$$k_2 = 2 - k_1 - k_3 = \frac{8}{9};$$

and finally

$$\int_{-\theta}^{\theta} P(x)dx = \frac{\theta}{9}\left\{5P\left(-\theta\sqrt{\frac{3}{5}}\right) + 8P(0) + 5P\left(\theta\sqrt{\frac{3}{5}}\right)\right\} \quad (36)$$

for any polynomial of fifth degree, since $i = 3 - 1 = 2$.

[4] *Goursat: Cours d'Analyse Mathématique*, vol. 1, p. 268; *Runge u. König: Numerisches Rechnen*, p. 275.

[5] *Crout: A Method for Deriving Formulas for the Approximate Calculation of Integrals*, this *Journal*, vol. 7, p. 126.

6. Conditioned Formulas with Invariancy of Order i. It was shown in Section 5 that the necessary and sufficient condition for a formula to have invariancy of order i is that the condition (34) be satisfied; however, since (34) involves $n - i - 1$ parameters $\beta_1, \beta_2, \cdots, \beta_{n-i-1}$, we can in general impose $n - i - 1$ conditions which fix these β's, the result being a definite formula of integration. The restriction in the choice of conditions is obviously that the system of relations for determining the β's be determinate.

As an illustration we may suppose that $n - i - 1$ *r's are fixed arbitrarily*; then the β's are determined so that these r's will be roots of (34). As an example let us consider the case where $n = 5$, $r_1 = -1$, $r_5 = 1$; then $n - i - 1 = 2$, or $i = 2$ and

$$\Pi = \frac{d^3}{du^3}\{(u^2 - 1)^3(u - \beta_1)(u - \beta_2)\},$$

from which it follows that

$$-\beta_1 = \beta_2 = 1,$$

and that $r_1, r_2, \cdots, r_5$ are the roots of

$$7u^5 - 10u^3 + 3u = 0.$$

But $r_1 = -1$, $r_5 = 1$; thus dividing by $u^2 - 1$ we see that the remaining three r's are roots of

$$7u^3 - 3u = 0;$$

hence $r_1 = -1$, $r_2 = -\sqrt{3/7}$, $r_3 = 0$, $r_4 = \sqrt{3/7}$, $r_5 = 1$; and finally from (8) we have

$$\int_{-\theta}^{\theta} P(x)dx = \frac{\theta}{90}\left\{9P(-\theta) + 49P\left(-\theta\sqrt{\frac{3}{7}}\right) + 64P(0) + 49P\left(\theta\sqrt{\frac{3}{7}}\right) + 9P(\theta)\right\}. \quad (37)$$

Since $i = 2$, (37) gives exact results when applied to any polynomial of seventh degree.

7. Approximation of $\int_{-\theta}^{\theta} f(x)dx$. The formulas which were derived in previous sections are such that exact results are obtained when they are applied to polynomials of degree smaller than some given number; hence it is evident that if such a formula is applied to an arbitrary function $f(x)$, the result will be the integral of a polynomial approximation of $f(x)$; thus

$$\int_{-\theta}^{\theta} f(x)dx = \theta \sum_{i=1}^{n} k_i y_i + \epsilon \tag{38}$$

where

$$\epsilon = \int_{-\theta}^{\theta} \{f(x) - P(x)\}dx.$$

By special methods [5] it can be shown that in the case of any particular formula there exists a relation of the form

$$|\epsilon| \leq \frac{\sigma_i \theta^{i+1} f_m^{\,i}}{\underline{i+1}} \qquad i \leq L, \tag{39}$$

where L and σ_i are constants, and $f_m^{\,i}$ is the greatest $|d^i f/dx^i|$ in the interval bounded by the two outermost points of the set $-\theta, \theta, r_1\theta, r_2\theta, \cdots, r_n\theta$.

A NEW TAUBERIAN THEOREM WITH APPLICATION TO THE SUMMABILITY OF FOURIER SERIES AND INTEGRALS

By S. B. Littauer

I. Recently N. Wiener [1] developed a general method for proving theorems termed Tauberian by Hardy and Littlewood. The theorems proved are of the form: if a function has a mean of a certain type at infinity, and if its growth is properly restricted, then it has many other means of specified types at infinity or at zero. By means of theorems of this type, N. Wiener is able to prove the O_A Tauberian theorems for Abel, Lambert, and Borel summation. The proof of the theorem on Lambert summation practically proves the classical theorem on the distribution of primes, and is the first method of proof that does not assume the validity of the theorem first. It was suggested to the present author by N. Wiener to develop other theorems of this character which might be applied to questions of the summability of Fourier series and integrals.

I am deeply indebted to Professor Wiener for his suggestion of the topic and for his guidance in the completion of this paper.

An interesting convergence criterion for Fourier series is Young's,[2] which is substantially: *for the Fourier series of the integral function $f(u)$ to converge to s for $u = x$, it is sufficient that*

$$\phi(t) \to 0 \tag{1.1}$$

and that

$$\int_0^t |d(u\phi(u))| = O(t) \tag{1.2}$$

for small t, where

$$\phi(t) = \tfrac{1}{2}\{f(x+t) + f(x-t) - 2s\}$$

[1] N. Wiener (1).
[2] Young (1).

This criterion was further developed by Young,[3] Pollard,[4] and Hardy[5] and Littlewood. Hardy and Littlewood's theorem is the most recent and most general. It is essentially: *the necessary and sufficient condition that the Fourier series of the integrable function $f(u)$ be summable $(C, -1+\delta)$ for any positive δ is that*

$$\phi_1(t) = o(t) \tag{1.3}$$

and

$$\int_0^t |d(u\phi(u)| = O(t), \quad 0 \leq t \leq \pi \tag{1.4}$$

where

$$\phi_1(t) = \int_0^t \phi(u)du \tag{1.5}$$

(1.2) and (1.4) suggest a general theorem of Tauberian type which in the present paper will be applied to proving the Hardy-Littlewood theorem using Riesz summability, instead of Cesaro. Although the present proof is much longer than Hardy-Littlewood's, it is believed that the general-theorem has more applications than are here given and that it will serve toward establishing the value of the general method developed by N. Wiener.

II(a). To set up the hypothesis of the general theorem, let

$$\phi(t) = \frac{1}{t}\int_0^t |ud\psi(u)| \tag{2_a.1}$$

This is permissible since (2_a.1) can be solved for $\psi(u)$. It follows directly that

$$d(t\phi(t)) = td\psi(t)$$

whence

$$\int_0^t |d(u\phi(u))| = O(t) = \int_0^t |ud\psi(u)|$$

which is equivalent to

$$\frac{1}{t}\int_0^t |ud\psi(u)| = O(1). \tag{2_a.2}$$

[3] Young (1.2).
[4] Pollard (1).
[5] Hardy and Littlewood (1).

Making the exponential transformation * $t = e^{\eta}$, $u = e^{\xi}$, we have,

$$e^{-\eta}\int_{-\infty}^{\eta} |e^{\xi}d\psi(e^{\xi})| = O(1) \leq \int_{\eta-1}^{\eta} e^{\xi-\eta}|d\psi(e^{\xi})| = O(1)$$

$$(2_a.3) \qquad \leq \int_{\eta-1}^{\eta} |d\psi(e^{\xi})| = \int_{\eta-1}^{\eta} |df(\xi)| = O(1)$$

independent of η. In further application of $(2_a.3)$ the limits used will be n and $n-1$ or $n+1$, where n represents any integer.

$(2_a.3)$ suggests setting up a general Tauberian theorem involving Stieltjes integrals, obeying $(2_a.3)$. With this as a starting point we may set down to be proved:

Theorem I:

If K_1 be a bounded measurable function, $O(1/\xi^2)$ at $\pm\infty$, K_2 a continuous function, also $O(1/\xi^2)$ at $\pm\infty$, and if the following conditions hold:

$$(2_a.4) \qquad \int_{n}^{n+1} |df(\xi)| = O(1),$$

$$(2_a.5) \qquad \lim_{\eta\to\pm\infty}\int_{-\infty}^{\infty} K_1(\xi)df(\xi-\eta) = A\int_{-\infty}^{\infty} K_1(\xi)d\xi,$$

$$(2_a.6) \qquad \int_{-\infty}^{\infty} K_1(\xi)df(\xi-\eta) = O(1), \quad [-\infty < \eta < \infty],$$

$$(2_a.7) \qquad \int_{-\infty}^{\infty} K_1(\xi)e^{iu\xi}d\xi \neq 0, \quad [-\infty < u < \infty],$$

then

$$(2_a.8) \qquad \lim_{\eta\to\pm\infty}\int_{-\infty}^{\infty} K_2(\xi)df(\xi-\eta) = A\int_{-\infty}^{\infty} K_2(\xi)d\xi.$$

* This transformation is equivalent to converting the average of a function about the origin to the average of the transformed function with a weighting factor, such that only the region of minus infinity of the function contributes to the average; and the average of a function at infinity becomes the average of the transformed function with a bounded weighting factor that has a specified rate of vanishing at $\pm\infty$, and such that only the region of $+\infty$ contributes to the result.

There are four essential steps to the proof, which I shall give in outline first before proceeding to the actual details, which though simple are somewhat lengthy. First any function *

$$(2_a.9) \qquad K(\xi) = \int_{-\infty}^{\infty} K_1(\xi - \zeta)R(\zeta)d\zeta$$

will be shown to satisfy $(2_a.8)$. Then it will be shown how to find a set of functions $K_a(\xi)$ which satisfy $(2_a.9)$ and in consequence $(2_a.8)$. Further it will be shown that any function $K(\xi)$ which is approximated to by a function of the class $K_a(\xi)$ in such a way that the sum over $(-\infty, \infty)$ of the absolute values of the maxima (per finite interval) of the differences between $K(\xi)$ and a $K_a(\xi)$ can be made arbitrarily small, satisfies $(2_a.8)$. Then an actual function of the class $K_a(\xi)$ satisfying $(2_a.8)$ will be formed, from which it can be readily shown that any continuous function satisfies $(2_a.8)$.

II(b) From the hypothesis,

$$(2_b.1) \qquad \int_{-\infty}^{\infty} K_1(\xi)df(\xi - \eta) - A\int_{-\infty}^{\infty} K_1(\xi)d\xi = F(\eta)$$

approaches 0 as $\eta \to \infty$. Consider then

$$(2_b.2) \qquad \left|\int_{-\infty}^{\infty} R(\zeta)F(\eta - \zeta)d\zeta\right| \leq \max_{(\zeta \leq B)} |F(\eta - \zeta)| \int_{-\infty}^{B} |R(\zeta)|d\zeta + \max|F(\eta - \zeta)| \int_{B}^{\infty} |R(\zeta)|d\zeta.$$

But by $(2_b.1)$ $\max|F(\eta - \zeta)|$ approaches 0 as $\eta \to \infty$ and since B is arbitrary, $\int_{B}^{\infty} |R(\zeta)|d\zeta$ approaches 0 for $R(\zeta)$ an absolutely integrable function. Hence,

$$(2_b.3) \qquad \overline{\lim_{\eta \to \infty}} \left|\int_{-\infty}^{\infty} R(\zeta)F(\eta - \zeta)d\zeta\right| = 0$$

In essentially the same manner it can be shown that $(2_b.3)$ is true for the limit as $\eta \to -\infty$.

* Where R is an absolutely integrable function.

$(2_b.3)$ is, however, equivalent to

$$(2_b.3_a) \quad \lim_{\eta \to \pm\infty} \int_{-\infty}^{\infty} R(\zeta)d\zeta \int_{-\infty}^{\infty} K_1(\xi)df(\xi - \eta + \zeta) = A \int_{-\infty}^{\infty} R(\zeta)d\zeta \int_{-\infty}^{\infty} K_1(\xi)d\xi.$$

If, then, it can be shown that

$$(2_b.3_b) \quad \lim_{\eta \to \pm\infty} \int_{-\infty}^{\infty} R(\zeta)d\zeta \int_{-\infty}^{\infty} K_1(\xi)df(\xi - \eta + \zeta) = \lim_{\eta \to \pm\infty} \int_{-\infty}^{\infty} df(\xi - \eta) \int_{-\infty}^{\infty} K_1(\xi - \zeta)R(\zeta)d\zeta$$

and that

$$(2_b.3_c) \quad A \int_{-\infty}^{\infty} R(\zeta)d\zeta \int_{-\infty}^{\infty} K_1(\xi)d\xi = A \int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} K_1(\xi - \zeta)R(\zeta)d\zeta$$

it follows that $K(\xi)$ which satisfies $(2_a.9)$ satisfies $(2_a.8)$.

Let the left side of $(2_b.3_a)$ be written as

$$(2_b.4) \quad \lim_{N \to \infty} \int_{-\infty}^{\infty} R(\zeta)d\zeta \sum_{-N}^{N} \lim_{\eta \to \pm\infty} \int_{\eta-\zeta}^{\eta+1-\zeta} K_1(\zeta)df(\xi - \eta + \zeta)$$

$$\leq \lim_{N \to \infty} \int_{-\infty}^{\infty} R(\zeta)d\zeta \sum_{-N}^{N} \max_{\eta-\zeta<\xi<\eta+1-\zeta} K_1(\xi) \lim_{\eta \to \pm\infty} \int_{\eta-\zeta}^{\eta+1-\zeta} df(\xi-\eta+\zeta).$$

The right side of $(2_b.4)$ (and hence the left side) will exist, since $R(\zeta)$ is absolutely integrable, and since

$$\lim_{\eta \to \pm\infty} \int_{\eta-\zeta}^{\eta+1-\zeta} df(\xi - \eta + \zeta) = O(1),$$

provided that

$$\lim_{N \to \infty} \sum_{-N}^{N} \max_{\eta-\zeta<\xi<\eta+1-\zeta} |K_1(\xi)| = O(1)$$

But for this to be true it is sufficient that $K_1(\xi)$ be bounded,

measurable, and $0(1/\xi^2)$ at $\pm\infty$. These are precisely the conditions put upon $K_1(\xi)$ in the hypothesis.

Further, by the facts and result of the preceding paragraph, we may write the left side of $(2_b.4)$ as,

$$\lim_{N\to\infty}\sum_{-N}^{N}\int_{-\infty}^{\infty}R(\zeta)d\zeta\lim_{\eta\to\pm\infty}\int_{\eta-\zeta}^{\eta+1-\zeta}K_1(\xi)df(\xi-\eta+\zeta)$$

which were it a finite sum, would permit of the inversion of the order of intergration, which can only be performed in the present case, if

$$(2_b.5)\qquad \lim_{N\to\infty}\sum_{-N+\zeta}^{N+\zeta}\lim_{\eta\to\pm\infty}\int_{\eta}^{\eta+1}df(\xi-\eta)\int_{-\infty}^{\infty}K_1(\xi-\zeta)R(\zeta)d\zeta$$

exists. $(2_b.5)$ will exist as double limit provided that

$$\lim_{N\to\infty}\sum_{-N}^{N}\max_{\eta<\xi-\zeta<\eta+1}|K_1(\xi-\zeta)|\int_{-\infty}^{\infty}|R(\zeta)|d\zeta$$

converges to a definite limit. This fact is established by the same argument as in the preceding paragraph, whence it has been shown that the order of integration of the left side of $(2_b.4)$ may be inverted. Therefore, $(2_b.3_b)$ holds.

By an analogous argument it can be shown that $(2_b.3_c)$ holds, whence $(2_b.3_a)$ may be converted to

$$(2_b.6)\qquad \lim_{\eta\to\pm\infty}\int_{-\infty}^{\infty}df(\xi-\eta)\int_{-\infty}^{\infty}K_1(\xi-\zeta)R(\zeta)d\zeta = A\int_{-\infty}^{\infty}d\xi\left(\int_{-\infty}^{\infty}K_1(\xi-\zeta)R(\zeta)d\zeta\right)$$

But $(2_b.6)$ is precisely the formulation of the fact that

$$K(\xi)=\int_{-\infty}^{\infty}K_1(\xi-\zeta)R(\zeta)d\zeta$$

where $R(\zeta)$ is an absolutely integrable function, satisfies $(2_a.8)$.

II(c) Suppose, now, we consider a class of functions,

$$K(\xi) = \int_{-\infty}^{\infty} K_1(\xi - \zeta)R(\zeta)d\zeta \tag{2$_c$.1}$$

and actually attempt to obtain such $K(\xi)$ by a general method, $K_1(\xi - \zeta)$ remaining a general and not a specific function satisfying the conditions stated in the hypothesis. By the Plancherel * theory of the Fourier transform the solution of (2$_c$.1) is

$$R(\zeta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{k(u)}{k_1(u)} e^{-iu\zeta}du \tag{2$_c$.2}$$

where

$$k(u) = \int_{-\infty}^{\infty} K(\xi)e^{iu\xi}d\xi$$

and

$$k_1(u) = \int_{-\infty}^{\infty} K_1(\xi)e^{iu\xi}d\xi$$

In order that the solution exist, $k_1(u)$ must be non-vanishing as prescribed in the hypothesis. Then $k(u)e^{-iu\xi}$ must be infinitely integrable with respect to u. It is sufficient that $k(u)$ be bounded between finite limits $-A \leqq u \leqq A$ and zero outside that range. That is, $K(\xi)$ must be a function such that its Fourier transform is bounded and vanishes for $|u| > A$. Further the negative Fourier transform of $k(u)$ must be absolutely integrable.

Take a $K^*(\xi)$, continuous and zero outside of $|\xi| = B$, finite. Its Fourier transform $k^*(u)$ is a bounded function of u. Now, take the negative Fourier transform of $k^*(u)$ between $-A$ and A. This is equivalent to mutilating $k^*(u)$ so as to be zero outside $|u| > A$. Call this last function, $K(\xi)$ the desired function, for the Fourier transform of $K(\xi)$ is $k^*(u)$ between $-A$ and A and zero elsewhere. This is sufficient to insure the existence of $R(\zeta)$.

* M. Plancherel (1).

Further $R(\zeta)$ will be absolutely integrable since $k(u)/k_1(u)$ is bounded and zero outside certain finite limits. The negative Fourier transform of this will be a bounded function, of order $1/\zeta^2$ at $\pm\infty$. This function, equivalent to $R(\zeta)$, with the above conditions is absolutely integrable.

In brief to find a function satisfying $2_a.8$, take the negative Fourier transform between finite limits of the Fourier transform of a continuous function which vanishes beyond certain finite values of its argument.*

II(d) Consider the class of functions $K_\alpha(\xi)$, satisfying $(2_a.8)$ and such that

$$(2_d.1) \qquad \lim_{\alpha\to\infty} \lim_{N\to\infty} \sum_{-N}^{N} \max_{\eta<\xi<\eta+1} |K_\alpha(\xi) - K(\xi)| = 0$$

Then

$$(2_d.2) \qquad \lim_{\alpha\to\infty} \lim_{\eta\to\pm\infty} \int_{-\infty}^{\infty} |df(\xi-\eta)|\,|K_\alpha(\xi) - K(\xi)|$$

$$= \lim_{\eta\to\pm\infty} \left\{ \lim_{\alpha\to\infty} \lim_{N\to\infty} \sum_{-N}^{N} \int_{\eta}^{\eta+1} |df(\xi-\eta)|\,|K_\alpha(\xi)-K(\xi)| \right\}$$

$$\leq \lim_{\eta\to\pm\infty} \left\{ \lim_{\alpha\to\infty} \lim_{N\to\infty} \sum_{-N}^{N} \left(\max_{\eta<\xi<\eta+1} |K_\alpha(\xi) - K(\xi)| \right.\right.$$

$$\left.\left. \times \int_{\eta}^{\eta+1} |df(\xi-\eta)| \right) \right\} = 0$$

since

$$\int_{\eta}^{\eta+1} |df(\xi-\eta)|$$

is bounded.

Hence

$$(2_d.3) \qquad \lim_{\alpha\to\infty} \lim_{\eta\to\pm\infty} \int_{-\infty}^{\infty} df(\xi-\eta)(K_\alpha(\xi) - K(\xi)) = 0$$

* This treatment is only heuristic. For a detailed proof, see N. Wiener (1).

and

$$(2_d.3_a) \quad \lim_{\alpha\to\infty}\lim_{\eta\to+\infty}\int_{-\infty}^{\infty} df(\xi-\eta)K_\alpha(\xi) = \lim_{\eta\to+\infty}\int_{-\infty}^{\infty} df(\xi-\eta)K(\xi)$$

But the left side of $(2_d.3_a)$ equals

$$A\left(\lim_{\alpha\to\infty}\int_{-\infty}^{\infty} K_\alpha(\xi)\,d\xi\right)$$

which from $(2_d.1)$ equals

$$A\int_{-\infty}^{\infty} K(\xi)d\xi$$

whence

$$(2_d.4) \quad \lim_{\eta\to+\infty}\int_{-\infty}^{\infty} df(\xi-\eta)K(\xi) = A\int_{-\infty}^{\infty} K(\xi)d\xi.$$

II(e) Now consider the following function

$$(2_e.1) \quad K(\xi) = \begin{cases} 0, & -\infty \le \xi < -\alpha \\ 1+\dfrac{\xi}{\alpha}, & -\alpha \le \xi \le 0 \\ 1-\dfrac{\xi}{\beta}, & 0 \le \xi \le \beta \\ 0, & \beta \le \xi \le \infty. \end{cases}$$

The Fourier transform of $K(\xi)$ is

$$(2_e.2) \quad \frac{1}{\sqrt{2\pi}}\int_{-\alpha}^{0}\left(1+\frac{\xi}{\alpha}\right)e^{iu\xi}d\xi + \frac{1}{\sqrt{2\pi}}\int_{0}^{\beta}\left(1-\frac{\xi}{\beta}\right)e^{iu\xi}d\xi$$
$$= \frac{1}{\sqrt{2\pi}}\cdot\frac{1}{u^2}\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)-\frac{e^{-iu\alpha}}{\alpha}-\frac{e^{iu\beta}}{\beta}\right]$$

By the argument in II(b), the negative Fourier transform of $(2_e.2)$ between finite limits will be a function satisfying (2.8). We wish to show that $(2_e.1)$ can be approximated to in the fashion of II(c) by such functions, as the finite limits are taken successively larger.

We have

$$(2_e.3)^* \quad \frac{1}{2\pi}\int_{-A}^{A}\frac{du}{u^2}\left(1-\frac{|u|}{A}\right)\left[\frac{e^{-iu\xi}}{\alpha}+\frac{e^{-iu\xi}}{\beta}-\frac{e^{-iu(\xi+\alpha)}}{\alpha}-\frac{e^{-iu(\xi-\beta)}}{\beta}\right]$$

$$=\frac{1}{2\pi}\int_{0}^{A}\frac{du}{u^2}\left(1-\frac{u}{A}\right)\left[\frac{e^{-iu\xi}}{\alpha}+\frac{e^{-iu\xi}}{\beta}-\frac{e^{-iu(\xi+\alpha)}}{\alpha}-\frac{e^{-iu(\xi-\beta)}}{\beta}\right]-\frac{1}{2\pi}\int_{A}^{0}\frac{du}{u^2}\left(1-\frac{u}{A}\right)\left[\frac{e^{-iu\xi}}{\alpha}+\frac{e^{-iu\xi}}{\beta}-\frac{e^{-iu(\xi+\alpha)}}{\alpha}-\frac{e^{-iu(\xi-\beta)}}{\beta}\right]$$

$$=\frac{1}{\pi}\int_{0}^{A}\frac{du}{u^2}\left(1-\frac{u}{A}\right)\left[\frac{e^{iu\xi}+e^{-iu\xi}}{\alpha}+\frac{e^{iu\xi}+e^{-iu\xi}}{\beta}-\frac{e^{iu(\xi+\alpha)}+e^{-iu(\xi+\alpha)}}{\alpha}-\frac{e^{iu(\xi-\beta)}+e^{-iu(\xi-\beta)}}{\beta}\right]$$

$$=\frac{1}{\pi}\int_{0}^{A}\frac{du}{u^2}\left(1-\frac{u}{A}\right)\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\cos u\xi-\frac{\cos u(\xi+\alpha)}{\alpha}-\frac{\cos u(\xi-\beta)}{\beta}\right]$$

$$=\frac{1}{\pi}\int_{0}^{A}\left(\frac{1}{u^2}-\frac{1}{uA}\right)\left[\frac{d\sin u\xi}{\alpha\xi}+\frac{d\sin u\xi}{\beta\xi}-\frac{d\sin(u(\xi+\alpha))}{\alpha(\xi+\alpha)}-\frac{d\sin u(\xi-\beta)}{\beta(\xi-\beta)}\right]$$

This last expression does satisfy $(2_e.8)$ from the argument of part II(b). It is therefore now necessary to estimate the difference between the limiting value of this function as A is permitted to grow large, and $K(\xi)$ given in $(2_e.1)$. This can be readily done if $K(\xi)$ be expressed as the negative Fourier transform of its own Fourier transform, and the difference taken before evaluating the integrals. The expression for the approximating function is gotten by going through the same process as was used in obtaining the last expression in $(2_e.3)$. Thus we

* The weighting function $1-(U)/A$ does not alter the character of the resulting function or its Fourier transform.

have for the difference

$$(2_c.4)\quad \frac{1}{\pi}\int_0^\infty \frac{1}{u^2}\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{d\sin u\xi}{\xi} - \frac{d\sin u(\xi+\alpha)}{\alpha(\xi+\alpha)} - \frac{d\sin u(\xi-\beta)}{\beta(\xi-\beta)}\right] - \frac{1}{\pi}\int_0^A \left(\frac{1}{u^2}-\frac{1}{uA}\right)\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{d\sin u\xi}{\xi} - \frac{d\sin u(\xi+\alpha)}{\alpha(\xi+\alpha)} - \frac{d\sin u(\xi-\beta)}{\beta(\xi-\beta)}\right]$$
$$= \frac{1}{\pi}\int_A^\infty \frac{1}{u^2}\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{d\sin u\xi}{\xi} - \frac{d\sin u(\xi+\alpha)}{\alpha(\xi+\alpha)} - \frac{d\sin u(\xi-\beta)}{\beta(\xi-\beta)}\right] + \frac{1}{\pi}\int_0^A \frac{1}{uA}\left[\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{d\sin u\xi}{\xi} - \frac{d\sin u(\xi+\alpha)}{\alpha(\xi+\alpha)} - \frac{d\sin u(\xi-\beta)}{\beta(\xi-\beta)}\right].$$

Upon applying to the above expression the transformations of variable, $u\xi = p$, $u(\xi+\alpha) = q$, and $u(\xi-\beta) = r$, and (since the definite integral is not altered by changing the symbol which represents the variable of integration) replacing that variable in each integral by z, one obtains, calling the approximating function $K_a(\xi)$

$$(2_c.5)\quad K(\xi) - K_a(\xi) = \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{1}{\pi}\int_{A\xi}^\infty \frac{\xi}{z^2}d\sin z + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{1}{\pi}\int_0^{A\xi}\frac{1}{Az}d\sin z$$
$$- \frac{1}{\alpha\pi}\left[\int_{A(\xi+\alpha)}^\infty \frac{(\xi+\alpha)}{z^2}d\sin z + \int_0^{A(\xi+\alpha)}\frac{1}{Az}d\sin z\right]$$
$$- \frac{1}{\beta\pi}\left[\int_{A(\xi-\beta)}^\infty \frac{(\xi-\beta)}{z^2}d\sin z + \int_0^{A(\xi-\beta)}\frac{1}{Az}d\sin z\right].$$

Integrating by parts,

$$\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{1}{\pi}\left[\int_{A\xi}^{\infty}\frac{\xi}{z^2}d\sin z+\int_0^{A\xi}\frac{1}{Az}d\sin z\right]$$
$$=\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{1}{\pi}\left\{\frac{\xi}{z^2}\sin z\Big|_{A\xi}^{\infty}+\frac{\sin z}{Az}\Big|_0^{A\xi}\right\}$$
$$+\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\frac{1}{\pi}\left[\int_{A\xi}^{\infty}\frac{\xi}{z^3}\sin z dz+\int_0^{A\xi}\frac{\sin z}{Az^2}dz\right];$$
$$-\frac{1}{\alpha\pi}\left[\int_{A(\xi+\alpha)}^{\infty}\frac{\xi+\alpha}{z^2}d\sin z+\int_0^{A(\xi+\alpha)}\frac{1}{Az}d\sin z\right]$$
$$=-\frac{1}{\alpha\pi}\left\{\frac{(\xi+\alpha)}{z^2}\sin z\Big|_{A(\xi+\alpha)}^{\infty}+\frac{\sin z}{Az}\Big|_0^{A(\xi+\alpha)}\right\}$$
$$-\frac{1}{\alpha\pi}\left[\int_{A(\xi+\alpha)}^{\infty}\frac{(\xi+\alpha)}{z^3}\sin z dz+\int_0^{A(\xi+\alpha)}\frac{\sin z}{Az^2}dz\right];$$
$$-\frac{1}{\beta\pi}\left[\int_{A(\xi-\beta)}^{\infty}\frac{\xi-\beta}{z^2}d\sin z+\int_0^{A(\xi-\beta)}\frac{1}{Az}d\sin z\right]$$
$$=-\frac{1}{\beta\pi}\left\{\frac{(\xi-\beta)}{z^2}\sin z\Big|_{A(\xi-\beta)}^{\infty}+\frac{\sin z}{Az}\Big|_0^{A(\xi-\beta)}\right\}$$
$$-\frac{1}{\beta\pi}\left[\int_{A(\xi-\beta)}^{\infty}\frac{(\xi-\beta)}{z^3}\sin z dz+\int_0^{A(\xi-\beta)}\frac{\sin z}{Az^2}dz\right].$$

Whence on substituting in the limits and combining similar integrals, the result is,

$$(2_a.6)\quad K(\xi)-K_a(\xi)=\frac{1}{\alpha\pi}\int_{A\xi}^{A(\xi+\alpha)}\left(\frac{\xi}{z^3}-\frac{1}{Az^2}\right)\sin z dz$$
$$-\frac{1}{\beta\pi}\int_{A(\xi-\beta)}^{A\xi}\left(\frac{\xi}{z^3}-\frac{1}{Az^2}\right)\sin z dz-\frac{1}{\pi}\int_{A(\xi+\alpha)}^{\infty}\frac{\sin z}{z^3}dz$$
$$+\frac{1}{\pi}\int_{A(\xi-\beta)}^{\infty}\frac{\sin z}{z^3}dz=0\left(\frac{1}{A^3\xi^2}\right).$$

With ξ fixed and $A\to\infty$ ($A\to\infty$ means $K_a(\xi)\to K_\infty(\xi)$).

$$\lim_{a\to\infty}\lim_{N\to\infty}\sum_{-N}^{N}\max_{\eta<\xi<\eta+1}|K(\xi)-K_a(\xi)|=0$$

whence $K(\xi)$ becomes a function of the class satisfying $(2_a.8)$.

Since any continuous function can be approximated to as closely as desired, by a series of functions of the type of $K(\xi)$, then any continuous function can be approximated to in the manner desired by a linear combination of functions which satisfy $(2_n.8)$. Obviously the proof for $K(\xi)$ holds for all functions of that type, and any linear combination of such functions satisfies $(2_n.8)$. Hence, the proof of the main theorem is complete.

III. It will be now shown that

Theorem II: *The necessary and sufficient condition that the Fourier series of the integrable function of $f(u)$, for $u = x$, be summable to s by any Riesz mean $(R, -1 + \delta)$, for any positive δ if*

$$\int_0^t |d(u\phi(u)| = O(t), \quad 0 \leq t \leq \pi, \tag{3.1}$$

is that

$$\phi_1(t) = \int_0^t \phi(u)du = O(t), \tag{3.2}$$

can be put into form so that it follows directly from the main theorem.

In

$$\frac{1}{t}\int_0^t \phi(u)du = o(1)$$

replace $\phi(u)$ by

$$\frac{}{u}\int_0^u |vd\psi(v)|$$

obtaining

$$\frac{1}{t}\int_0^t \frac{du}{u}\int_0^u vd\psi(v) = o(1)$$

which on inverting the order of integration becomes

$$\frac{1}{t}\int_0^t d\psi(v)\int_v^t \frac{v}{u}du = \frac{1}{t}\int_0^t d\psi(v)v\log\left(\frac{t}{v}\right) = o(1).$$

On applying the transformation $t = e^{-\eta}$, $v = e^{\xi}$ the above expression becomes

$$-\int_{-\infty}^{\eta} d\psi(e^{\xi})e^{\xi+\eta}(\eta+\zeta) = o(1)$$

which is now close to the form of the hypothesis of the main theorem. Again replacing ζ by $\xi - \eta$ we have

$$\lim_{\eta\to\infty} -\int_{-\infty}^{0} d\psi(e^{\xi-\eta})[e^{\xi}\xi] = o(1) \tag{3.3}$$

as the final form of (3.2). This it will be observed is exactly analogous to $(2_a.5)$ where $K_1(\xi)$ is now $[-e^{\xi}\xi]$ for $-\infty \leq \xi \leq 0$ and 0 for $0 < \xi \leq \infty$.

The expression for the limit of the *nth partial* remainder as $n \to \infty$ of the kth * order Riesz mean of the Fourier series of $f(u)$ for $u = x$ is

$$\lim_{w\to\infty} \int_0^{\infty} t^{-(1+k)}C_{1+k}(t)\phi(t/w)dt \dagger \tag{3.4}$$

where

$$\phi(t) = \tfrac{1}{2}\{f(x+t) + f(x-t) - 2s\}$$

and

$$C_{1+k}(t)t^{-(1+k)} = \frac{1}{\Gamma(1+k)}\int_0^1 (1-\lambda)^k \cos \lambda t d\lambda. \ddagger$$

That the Fourier series be convergent as it is here summed, (3.4) must equal 0.

Again (3.4) must be transformed so as to be analogous to $(2_a.6)$. First since $\Gamma(1+k)$ contains no zeros it may be omitted from $C_{1+k}(t)t^{-(1+k)}$ without affecting the proof that (3.4) = 0. On making there the replacement for $C_{1+k}(t)t^{-(1+k)}$ and for $\phi(t/w)$ using $(2_a.1)$, letting $t = t/w$, (3.4) becomes,

$$\lim_{w\to\infty} \int_0^{\infty} \frac{wdt}{t}\int_0^t ud\psi(u)\int_0^1 (1-\lambda)^k \cos \lambda wtd\lambda \tag{3.5}$$

Now, since $\frac{1}{t}\int_0^t |ud\psi(u)|$ is bounded, and since $\int_0^1 (1-\lambda)^k$

* $k = -1 + \delta$ for a positive δ.
† Hobson (1).
‡ Jacob (1).

$\cos \lambda wtd\lambda$ * is absolutely integrable in ωt, (3.5) exists as the limit of a sum of all positive terms, independent of ω, whence the order of integration in (3.5) may inverted, giving

$$(3.6) \qquad \lim_{\omega \to \infty} \omega \int_0^\infty u d\psi(u) \int_u^\infty \frac{dt}{t} \int_0^1 (1-\lambda)^k \cos \lambda \omega t d\lambda$$

Applying the transformation $u = e^{\xi - \eta}$, $\omega = e^\eta$ (3.6) becomes,

$$(3.7) \qquad \lim_{\eta \to \infty} \int_{-\infty}^\infty d\psi(e^{\xi-\eta}) e^\xi \int_{\xi-\eta}^\infty \frac{dt}{t} \int_0^1 (1-\lambda)^k \cos \lambda e^\eta t d\lambda,$$

which is the form into which it is desired to cast (3.4).

Then the proof of theorem II becomes equivalent to proving that *when* (3.1) *holds*: *if* (3.3) *is true*, (3.7) = 0, *and if* (3.7) = 0, (3.3) *is true*. The proof follows immediately from the main theorem upon showing that the conditions of that theorem are fulfilled.

Consider first the kernel of (3.3)

$$K_1(\xi) = \begin{cases} [-e^k \xi] & \text{for} \quad -\infty \leq \xi \leq 0 \\ 0 & \text{for} \quad 0 < \xi \leq \infty \end{cases}$$

At $\xi = \pm \infty$ it is obviously $0(1/\xi^2)$. Further $K_1(\xi)$ is continuous.

Since $\int_\eta^{\eta+1} |d\psi(e^{\xi-\eta})|$ is bounded, and since $K_1(\xi)$ is bounded and $0(1/\xi^2)$ at $\pm \infty$, independent of η,

$$(3.8) \qquad \int_{-\infty}^0 d\psi(e^{\xi-\eta}) [-e^k \xi] = O(1), \quad [-\infty < \eta < \infty]$$

Further

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 - e^k \xi e^{iu\xi} d\xi$$

on letting $e^k = x$ reduces to

$$(3.9) \qquad -\frac{1}{\sqrt{2\pi}} \int_0^1 x^{iu} \log x dx = \frac{1}{(iu+1)^2}$$

* Jacob (1).

which has no zeros for $-\infty < u < \infty$. Thus the $K_1(\xi)$ of Theorem II satisfies all the conditions put upon it in the main theorem.

Consider now, the kernel of (3.7)

$$K_2(\xi) = e^{\xi} \int_{e^{\xi-\eta}}^{\infty} \frac{dt}{t} \int_0^1 (1-\lambda)^k \cos \lambda e^{\eta} t d\lambda. \tag{3.10}$$

First

$$\int_0^1 (1-\lambda)^k \cos \lambda e^{\eta} t d\lambda = 0(t^{-1-k}).^*$$

Hence at $\xi = +\infty$

$$K_2(\xi) = O\left(e^{\xi} \frac{1}{e^{\xi(1+\delta)}}\right), \quad \delta > 0 \tag{3.11}$$

$$= O\left(\frac{1}{e^{\xi\delta}}\right) < O\left(\frac{1}{\xi^2}\right)$$

For $\xi = -\infty$, consider

$$K_2(\xi) = e^{\xi} \int_0^{\infty} O(t^{-2-k}) dt \leq e^{\xi} O(|\xi|^{1+\delta})$$

which as $\xi \to -\infty$

$$< O\left(\frac{1}{\xi^p}\right), \quad p < n \tag{3.11$_a$}$$

where n is any finite power. Thus $K_2(\xi)$ is of the proper order of magnitude at $\pm\infty$.

It was shown, above that

$$\int_0^{\infty} \omega \frac{dt}{t} \int_0^t u d\psi(u) \int_0^1 (1-\lambda)^k \cos \lambda \omega t d\lambda$$

is bounded independent of ω. Whence

$$\int_{-\infty}^{\infty} d\psi(e^{\xi-\eta}) e^{\xi} \int_{e^{\xi-\eta}}^{\infty} \frac{dt}{t} \int_0^t (1-\lambda)^k \cos \lambda e^{\eta} t d\lambda \tag{3.12}$$

is bounded $-\infty < \eta < \infty$.

* M. Jacob (1).

Now to evaluate

$$k(u) = \int_{-\infty}^{\infty} \left(e^{\xi} \int_{e^{\xi}-\eta}^{\infty} \frac{dt}{t} \int_0^1 (1-\lambda)^k \cos \lambda e^{\eta} t d\lambda \right) e^{i\xi u} d\xi$$

By letting $t = e^{\eta}\tau$

$$k(u) = \int_{-\infty}^{\infty} e^{\xi + i\xi u} \int_{e^{\xi}}^{\infty} \frac{d\tau}{\tau} \int_0^1 (1-\lambda)^k \cos \lambda \tau d\lambda$$

which for $e^{\xi} = x$ becomes,

$$\begin{aligned} (3.13) \qquad k(u) &= \int_0^{\infty} x^{iu} dx \int_x^{\infty} \frac{d\tau}{\tau} \int_0^1 (1-\lambda)^k \cos \lambda \tau d\lambda \\ &= \int_0^{\infty} \frac{d\tau}{\tau} \int_0^{\tau} x^{iu} dx \int_0^1 (1-\lambda)^k \cos \lambda \tau d\lambda \\ &= \frac{\Gamma(1+k)}{iu+1} \int_0^{\infty} c_{1+k}(\tau) \tau^{-1-k+iu} d\tau \\ &= \frac{\pi}{2} \left(\frac{1}{iu+1} \right) \frac{\Gamma(1+k)}{\Gamma(1+k-iu)} \frac{1}{\cos \dfrac{iu\pi}{2}} \\ &\neq 0, \quad [-\infty < u < \infty]. \end{aligned}$$

It remains to show that $K_2(\xi)$ is continuous. This, follows directly from the fact that $\int_0^1 (1-\lambda)^k \cos \lambda e^{\eta} t d\lambda$ is a continuous function of t.

Thus, it has been shown that $K_2(\xi)$ of theorem II satisfies all the conditions put upon it by the main theorem. Theorem II then is completely a special case of theorem I, whence theorem II is proved.

This proof of theorem II, while actually a little longer than the Hardy-Littlewood proof, seems to be essentially simpler. The conversion of theorem II into a form which is a particular case of the general theorem is accomplished by direct and elementary means. The Hardy-Littlewood proof has recourse to four lemmas, and some intricate and none too obvious analysis,

and may not be applied in method to other theorems. Of course the proof for Riesz summability is simpler than that for the Cesaro sums.

It is of further interest to note that the present method of proof throws light on why the Young criterion is so effective. For the Tauberian theorem shows that certain averages of a function which is only restricted by the fact that the integral of its differential per finite interval be bounded, exists and are equal. While the Young criterion says that if $f(x)$ be of the type given above then the Fourier series of the function is convergent. Then the added Hardy-Littlewood condition further asserts that any Riesz (or Cesaro * as will be shown later) sum of order greater than -1 is the average of the given function multiplied by a weighting factor which is continuous, whence by the general theorem of this paper the average exists, and is the same for any order.

IV. It is well to note that a theorem II is proved identically as above for the summability of the Fourier integral. (For the integral, of course, Riesz and Cesaro sums of any order are identical.) The identity of proof follows from the fact that the expression for the limit of the partial remainders of Riesz-Cesaro sum of the Fourier integral is identical with (3.7), as will be directly shown. This remainder is,

$$\lim_{\omega\to\infty}\int_0^{\omega u}\phi\left(\frac{t}{\omega}\right)dt\int_0^1(1-u)^k\cos utdu\ \dagger \tag{4.1}$$

which is equivalent to, omitting a factor $\Gamma(1+k)$

$$\lim_{\omega\to\infty}\int_0^{\infty}\phi(t/\omega)C_{1+k}(u)u^{-(1+k)}dt \tag{4.2}$$

which is precisely (3.4) from which (3.7) was derived.

It remains to consider the proof of the Hardy-Littlewood theorem precisely for negative order Cesaro summability of the

* For integral only.
† Hobson (1), page 739.

series. For positive orders, since the Riesz and Cesaro sums of series are equivalent, it is indifferent for which sum a proof be given, it holds for the other. As far as the author knows, the literature does not contain a proof of the equivalence of Riesz and Cesaro sums of negative orders. And as yet the author has been unable to cast the Cesaro sum of a series into a form to which the present method of proof will be applicable without the additional requirement of evaluating a remainder. Hence further work on theorem II will be along the lines of simplifying the integral representing a Cesaro sum of a Fourier series, or of proving the equivalence of Cesaro and Riesz sums for negative orders.

V. In conclusion may it be said that the method of proving Tauberian theorems developed by Norbert Wiener is quite powerful and general, and promises to have very wide application. That this method is effective in handling certain types of summability. That Riesz summability although of later discovery than Cesaro summability, is simpler and closer to the essence of the whole question of summability than is its historical predecessor.

References

G. H. Hardy and J. E. Littlewood, (1) On Young's Convergence Criterion for Fourier Series, *Proc. Lond. Math. Soc.*, series 2, vol. 28 (1928), pp. 301–311.

Hobson, (1) The Theory of Functions of a Real Variable and Fourier Series, vol. II.

M. Jacob, (1) Über die Summierbarkeit von Fourierschen Reihen und Integralen, *Math. Zeit.*, vol. 29 (1928), pp. 20–33.

J. E. Littlewood, see G. H. Hardy.

M. Plancherel, (1) Contribution a l'etude de la representation d'une function arbitraire par des integrales definies. *Rendiconti di Palermo*, vol. 30 (1910), pp. 289–335.

S. Pollard, (1) On the Criteria for the Convergence of a Fourier Series, *Jour. Lond. Math. Soc.*, vol. 2 (1927), pp. 255–262.

N. Wiener, (1) A New Method of Tauberian Theorems, *Jour. of Math. and Phys., M.I.T.*, vol. VII (1928), pp. 161–184.

J. W. Young, (1) On the Convergence of the Derived Series of Fourier Series, *Proc. Lond. Math. Soc.*, (2) 17 (1916), pp. 195–236.

(2) Sur les conditions de convergence des series de Fourier, *Comptes rendus*, vol. 26, Dec., 1918.

(3) On Infinite Integrals Involving a Generalization of the Sine and Cosine Functions, *Quarterly Jour. of Math.*, vol. 43 (1911), pp. 161–177.

THE OPERATORS OF QUANTUM MECHANICS[1]

By Philip Franklin

1. Introduction. In the modern formulations of quantum mechanics as a characteristic value (Eigenwert) problem[2] for an infinite algebraic or differential system, one of the important questions which presents itself is the determination of the appropriate operator for a specific problem whose classical formulation is known. Each of the three theories quoted proceeds on the assumption that there is an operator corresponding to each given classical Hamiltonian function and that it is uniquely determined by the assumptions made. For the simplest cases the operator is explicitly constructed, and the impression is given that a general method is outlined. It is the purpose of this note to show that when the algorithms used are applied to the general case none of them is completely satisfactory. Only one of the methods, that of Schrödinger, is invariant under point transformations and this method only applies to Hamiltonian functions quadratic in the momenta. The other two methods are inconsistent with each other, as well as with this one. They are untenable, as they fail to give correct results for some of the cases for which the theory has been checked experimentally. The argument we give to show that they do not transform properly, proves that it is *impossible* to construct a satisfactory *linear* correspondence of operators (K) and Hamiltonian functions (H), i. e. such that if

[1] Presented to the American Mathematical Society, February, 1929.

[2] (a) W. Heisenberg, *Zeitscher. f. Phys.*, vol. 33 (1925), p. 879.
M. Born, P. Jordan, *Zeitschr. f. Phys.*, vol. 34 (1925), p. 858.
M. Born, W. Heisenberg, P. Jordan, *Zeitschr. f. Phys.*, vol. 35 (1926), p. 557.

(b) E. Schrödinger, *Ann. d. Phys.*, vol. 79, p. 361, p. 489, p. 734; vol. 80, p. 437; vol. 81, p. 109 (1926).
These papers are included in the collection: E. Schrödinger, *Abhandlungen zur Wellenmechanik*, Leipzig, 1927.

(c) H. Weyl, *Zeitschr. f. Phys.*, vol. 46 (1927), p. 1.
H. Weyl, *Gruppentheorie und Quantenmechanik*, Leipzig, 1928.

$$K_1 \sim H_1 \quad \text{and} \quad K_2 \sim H_2,$$

then

$$K_1 + K_2 \sim H_1 + H_2.$$

2. The ambiguity in the assumption of formal equivalence. Before describing in detail the various methods to be compared, we shall call attention to some properties of the formal equivalence between the operator and the classical function which is assumed by all of the writers at some stage of their argument. To be precise, they demand that if the operator, written in a *suitable* (well-ordered) form as

$$K(P_i, Q_i),$$

is equivalent to the classical Hamiltonian function

$$H(p_i, q_i),$$

then, when the operators are replaced by the quantities,

$$K(p_i, q_i) = H(p_i, q_i),$$

identically. Since for the operators

$$P_iQ_i - Q_iP_i = \frac{h}{2\pi i}, \tag{1}$$

while for the quantities

$$p_iq_i - q_ip_i = 0,$$

it is obvious that the condition that K be well-ordered is vital. It is also fairly evident, and has been pointed out by numerous writers, including Schrödinger himself [3] that K is not uniquely determined by H. It is a little less obvious that H is not uniquely determined by K, and apparently not generally noticed that the mere assumption of formal equivalence in the *absence* of other restrictions, places *no limitations whatever* on the correspondence. We shall accordingly explicitly demonstrate that

Any operator $F(P_i, Q_i)$ may be so ordered that it is formally equivalent to any function $f(p_i, q_i)$.

[3] E. Schrödinger, *l. c.* [2](b), vol. 79, p. 747.

Let $G(P_i, Q_i)$ be a particular operator corresponding to $f(p_i, q_i)$, and write:

$$\begin{aligned} F(P_i, Q_i) &= G(P_i, Q_i) + F(P_i, Q_i) - G(P_i, Q_i) \\ &= G(P_i, Q_i) + \frac{2\pi i}{h}[P_1Q_1 - Q_1P_1][F(P_i, Q_i) - G(P_i, Q_i)] \\ &= K(P_i, Q_i). \end{aligned}$$

We may regard this as the required ordered form for $F(P_i, Q_i)$, since, from the definition of $G(P_i, Q_i)$, we shall have:

$$K(p_i, q_i) = G(p_i, q_i) = f(p_i, q_i).$$

3. The correspondence of Born, Heisenberg, Jordan. In the first presentation of the matrix calculus,[4] attention is confined to Hamiltonian functions which are power series in the p_i and q_i. In addition to formal equivalence, it is demanded that the operator H satisfy several conditions when in the well-ordered form. The partial derivatives must satisfy:

$$\begin{aligned} HQ_i - Q_iH &= \frac{h}{2\pi i}\frac{\partial H}{\partial P}, \\ HP_i - P_iH &= -\frac{h}{2\pi i}\frac{\partial H}{\partial Q}, \end{aligned} \tag{3}$$

and these derivatives must be formally equivalent with the classical ones so as to yield " the same " Hamiltonian equations. Also, the operator must in matrix form be a diagonal matrix. We note that these are rather restrictions on H than on the correspondence, since if $F(P_i, Q_i)$ and $G(P_i, Q_i)$ are diagonal matrices satisfying (3), and we define $K(P_i, Q_i)$ by (2) above, this operator satisfies all the conditions as to form, and we have:

$$\begin{aligned} \frac{\partial K}{\partial P_s} &= \frac{\partial G}{\partial P_s} + \frac{2\pi i}{h}(P_1Q_1 - Q_1P_1)\left(\frac{\partial F}{\partial P_s} - \frac{\partial G}{\partial P_s}\right), \\ \frac{\partial K}{\partial Q_s} &= \frac{\partial G}{\partial Q_s} + \frac{2\pi i}{h}(P_1Q_1 - Q_1P_1)\left(\frac{\partial F}{\partial Q_s} - \frac{\partial G}{\partial Qs}\right), \end{aligned} \tag{4}$$

[4] M. Born and P. Jordan, *l. c.* [2](a), vol. 34, pp. 871–874.

so that we have the required formal equivalence:

$$\begin{aligned}(5)\qquad \frac{\partial K(p_i, q_i)}{\partial P_s} &= \frac{\partial G(p_i, q_i)}{\partial P_s} = \frac{\partial f(p_i, q_i)}{\partial p_s},\\ \frac{\partial K(p_i, q_i)}{\partial Q_s} &= \frac{\partial G(p_i, q_i)}{\partial Q_s} = \frac{\partial f(p_i, q_i)}{\partial q_s}.\end{aligned}$$

These considerations make it plausible that the assumptions just stated are not sufficient to define a unique correspondence. The authors under discussion obtain an operator for $p^m q^n$, possibly by using certain unstated assumptions. For example, the requirement that the operator should be a homogeneous polynomial of degree $m + n$ would rule out the device used in (2). Their result is [5]

$$(6)\qquad p^m q^n \sim \frac{1}{m+1}\sum_{r=0}^{m} P^r Q^n P^{m-r}.$$

This certainly satisfies all the assumptions we have quoted. As the linear property is assumed, the operator for any power series in one pair of variables can be derived from this result.

For later use we quote here two formulas [6] derived as consequences of (1):

$$\begin{aligned}(7)\qquad P^m Q^n - Q^n P^m &= m\frac{h}{2\pi i}\sum_{s=0}^{n-1} Q^{n-1-s}P^{m-1}Q^s\\ &= n\frac{h}{2\pi i}\sum_{s=0}^{n-1} P^{m-1-s}Q^{n-1}P^s.\end{aligned}$$

4. The correspondence of Weyl. The definition of the correspondence given by Weyl is more comprehensive.[6] He writes

$$(8)\qquad f(p, q) = \int_{-\infty}^{\infty}\int e^{i(p\sigma+q\tau)}\xi(\sigma, \tau)d\sigma d\tau$$

as the well-ordered form for a given function, with

$$\begin{aligned}(9)\qquad e^{i(p\sigma+q\tau)}G(q) &= e^{i(P\sigma+Q\tau)}G(q)\\ &= e^{(ih\sigma\tau)/4\pi}e^{i\tau Q}G\left(q + \frac{h\sigma}{2\pi}\right).\end{aligned}$$

[5] M. Born, W. Heisenberg, P. Jordan, *l. c.*, [2](b), p. 562, footnote.
[6] H. Weyl, *l. c.*, [2](c), vol. 46, pp. 27, 28, or his book. 205.

The essential idea is that (9) holds, and that the correspondence is *linear*. Consequently, if (8), as written, diverges, it is replaced by some other linear operation. For example, it may be made to converge [7] by putting $e^{-a|q|}f(p, q)$ in place of $f(p, q)$ and taking the limit of the transform as a approaches 0.

To compare this method with (6), we must apply it to $p^m q^n$. Here (8) diverges if applied directly, but is applicable to $e^{-ap-bq}p^m q^n$. We find:

$$(10)\quad e^{-ap-bq}p^m q^n = \frac{m!\,n!}{\pi^2}\int_{-\infty}^{\infty}\int e^{i(p\sigma+q\tau)}\frac{d\sigma}{(a+i\sigma)^{m+1}}\frac{d\tau}{(b+i\tau)^{n+1}}.$$

We obtain the equivalent operator from this by using (9).

$$(11)\quad e^{-ap-bq}p^m q^n \sim K_{ab},$$

$$(12)\quad K_{ab}G(q) = (-1)^m\left(\frac{d}{da}\right)^m e^{-b(q+(iha/4\pi))}\left(q+\frac{iha}{4\pi}\right)^n G\left(q+\frac{iha}{2\pi}\right).$$

On taking the limit as a and b each approach zero, we get in place of these the relations:

$$(13)\quad p^m q^n \sim K,$$

$$(14)\quad \begin{aligned} K &= \sum_{r=0}^{\min(m,n)} {}_mC_r\frac{n(n-1)\cdots(n-r+1)}{2^r}\left(\frac{h}{2\pi i}\right)^m q^{n-r}\left(\frac{\partial}{\partial q}\right)^{m-r} \\ &= \frac{1}{2^m}\sum_{r=0}^{m} {}_mC_r P^r Q^n P^{m-r}. \end{aligned}$$

The last form is obtained from (7), and the relation:

$$(15)\quad P = \frac{h}{2\pi i}\frac{\partial}{\partial q}.$$

5. The correspondence of Schrödinger. If we take the point of view of Schrödinger,[8] we form the operator for a given energy function by using a variation principle. If in the space of the q_j,

$$ds^2 = \sum g_{jk}dq_j dq_k, \quad g = |g_{jk}|,$$

[7] H. Weyl, *l. c.* [2](c), vol. 46, p. 39.
[8] E. Schrödinger, *l. c.*, [2](c), vol. 79, p. 748, or p. 376.

so that the volume element is

$$\sqrt{g}dq_1dq_2 \cdots dq_n,$$

the operator corresponding to any classical Hamiltonian function

$$T(p_j, q_j) + V(q_j),$$

where T is *homogeneous and quadratic in the* p_j, is

$$\frac{1}{2\sqrt{g}} \sum_k \frac{\partial}{\partial q_k} \left[\sqrt{g} \frac{\partial T\left(\frac{h}{2\pi i} \frac{\partial}{\partial q_j}, q_j\right)}{\partial p_k} \right] + V(q_j). \tag{16}$$

This is obtained by setting

$$\delta \int \cdots \int \left\{ - T\left(\frac{h}{2\pi i} \frac{\partial \psi}{\partial q_j}, q_j\right) + \psi^2 V(q_j) \right\} \sqrt{g}dq_1dq_2 \cdots dq_n = 0,$$

with the normalizing condition

$$\int \cdots \int \psi^2 \sqrt{g}dq_1dq_2 \cdots dq_n = 1.$$

If we insert a group of linear terms in the Hamiltonian function, $p_kL_k(q_j)$ it is not precisely clear how we should apply the variational method to them, but it is natural to include them in the integral, multiplied by $c\psi$, and this would give:

$$p_kL_k \sim c \left\{ - \frac{1}{2\sqrt{g}} \frac{h}{2\pi i} \frac{\partial}{\partial q_k} [\sqrt{g} L_k(q_j)] + \frac{h}{4\pi i} L_k(q_j) \frac{\partial}{\partial q_k} \right\}. \tag{17}$$

The constant c would have to be determined by further properties of problems in which the linear terms occurred.

When we attempt to extend the method to terms of higher degree in p_j, we find that the operator resulting from obvious natural extensions is not a *linear* operator, and hence the variational method seems to be limited to the quadratic case. While non-quadratic forms in the p_j do not occur directly, they appear when we change coördinates, e. g. by the contact transformation

$$\bar{p}_i = q_i, \quad \bar{q}_i = -p_i.$$

Furthermore, such cases are of interest in comparing the various general definitions of the correspondence of operators and functions.

6. Comparison of the methods. We first note that the operator for $p^m q^n$ given by Born, Heisenberg, Jordan, (6) is only consistent with that given by Weyl (14) in the simplest cases. In fact, these operators are identical for $m = 0, 1$ and for all n, as well as for $n = 0, 1$ for all m, as we see by applying (1). They *differ*, in general, for all other values. For example, for $m = 2$, we find by a double application of (7) that the difference of the two operators is:

$$(18)\quad \tfrac{1}{3}(P^2Q^n + PQ^nP + Q^nP^2) - \tfrac{1}{4}(P^2Q^n + 2PQ^nP + Q^nP^2) = -\frac{n(n-1)}{12}\frac{h^2}{4\pi^2}Q^{n-2},$$

which is a non-zero operator for $n \geq 2$.

To determine which of these methods is preferable, we shall extend their scope. We recall that each was derived from the assumption that the correspondence to be found had a linear character. Thus, if $f(q)$ is a polynomial, or more generally such that its operator may be found from its power series expansion, we find from (6)

$$(19)\qquad p^m \cdot f(q) \sim \frac{1}{m+1}\sum_{r=0}^{m} P^r f(Q) P^{m-r}.$$

In particular:

$$(20)\qquad p_1 \cdot f(q_j) \sim \tfrac{1}{2}(fP_1 + P_1 f),$$

$$(21)\qquad p_1^2 \cdot f(q_j) \sim \tfrac{1}{3}(fP_1^2 + P_1 fP_1 + P_1^2 f),$$

Since P_1 commutes with P_2 and Q_2, we should expect from (20) that

$$(22)\quad p_1 p_2 \cdot f(q_j) \sim \tfrac{1}{4}(fP_1P_2 + P_1 fP_2 + P_2 fP_1 + P_1P_2 f).$$

The last three equations combine to give as the operator for

$$T = \sum g_{jk} q_j' q_k' = \sum g^{jk} p_j p_k, \tag{23}$$

$$\sum \tfrac{1}{4}(1 - \delta_b^a)(g^{ab}P_aP_b + P_ag^{ab}P_b + P_bg^{ab}P_a + P_aP_bg^{ab}) \tag{23}$$

$$+ \sum \tfrac{1}{3}(g^{aa}P_a^2 + P_ag^{aa}P_a + P_a^2g^{aa}), \tag{24}$$

or,

$$\sum P_ag^{ab}P_b - \frac{h^2}{4\pi^2}\sum\left\{\frac{1}{4}(1 - \delta_b^a)\frac{\partial^2 g^{ab}}{\partial q_a \partial q_b} + \frac{1}{3}\frac{\partial^2 g^{aa}}{\partial q_a^2}\right\}. \tag{25}$$

We may treat (14) similarly, and find:

$$p^m f(q) \sim \frac{1}{2^m}\sum_{r=0}^{m} {}^mC_r P^r f(Q) P^{m-r}, \tag{26}$$

which leads to (20), and hence (22) as before, but gives in place of (21)

$$p_1^2 \cdot f(q_j) \sim \tfrac{1}{4}(fP_1^2 + 2P_1fP_1 + P_1^2f). \tag{27}$$

This, with (20) and (22), gives as the operator for the T of (23):

$$\sum \tfrac{1}{4}(g^{ab}P_aP_b + P_ag^{ab}P_b + P_bg^{ab}P_a + P_aP_bg^{ab}) \tag{28}$$

or,

$$\sum P_ag^{ab}P_b - \frac{h^2}{4\pi^2}\sum\frac{\partial^2 g^{ab}}{\partial q_a \partial q_b}. \tag{29}$$

The operator for the T of (23) as given by the method of Schrödinger, (16) is:

$$\frac{1}{\sqrt{g}} P_a \sqrt{g} g^{ab} P_b, \tag{30}$$

which is to be compared with (25) and (29).

In the case of constant g^{jk}, substantially the case in which the q_j are Cartesian coördinates, the formulas all agree, but when the g^{jk} are not constant, but vary with the q_j, equations (25), (29) and (30) will in general give three different results. Even when g is constant, so that (30) agrees with the first term of (25) and (29), the remaining terms will only vanish when the

g^{ik} satisfy further conditions, somewhat simpler for (29) but not easy to interpret in either case.

As (30), for fairly general g^{ik}, may be obtained from

$$\sum p_k^2 \sim \sum P_k^2, \tag{31}$$

by a point transformation, and, besides, has been used directly in some of the calculations which check the theory, the argument just given shows that the first two methods are untenable.

7. The impossibility of a linear correspondence. We have seen that two different methods of correspondence, each of which assumes:

1. That to each ordinary function $T(p_j, q_j)$ there is a *uniquely* determined operator, $K(P_j, Q_j)$;

2. That to a linear combination of two functions, $c_1T_1(p_j, q_j) + c_2T_2(p_j, q_j)$ there corresponds the same combination of the operators: $c_1T_1(P_j, Q_j) + c_2T_2(P_j, Q_j)$;

lead to incorrect results. We shall now show that:

No correspondence satisfying (1) *and* (2) *can be consistent with* (30), and hence these assumptions lead to incorrect results.

For, consider the particular function:

$$Ap_1^2 + Be^{-2f}p_2^2 + p_3^2 = T(p_j, q_j), \tag{32}$$

where A and B are constants, and f is a function of q_1 only. By (30), the corresponding operator should be:

$$A\left(P_1^2 + \frac{df}{dq_1}P_1\right) + Be^{-2f}P_2^2 + P_3^2. \tag{33}$$

But, if the assumptions (1) and (2) held, for varying A, this would necessitate

$$p_1^2 \sim P_1^2 + \frac{df}{dq_1}P_1, \tag{34}$$

so that the operator for p_1^2 could not be uniquely determined.

Professor Weyl was lead to his method, (8), (9) by the fact that this correspondence is invariant for an infinitesimal contact

transformation *starting out from the Cartesian case*, (31). If the correspondence remained invariant under successive applications of the infinitesimal transformation, and hence under finite transformations, the definition would hold for all coördinates obtained from Cartesian ones by contact transformations of a special type, namely those continuously deformable to the identity without meeting any singularity. The example we have just given does not preclude this, as it is not reducible to the Euclidean element for a general choice of f. For a special choice, it leads to the expression for the arc element in polar coördinates (for which the expression (29) breaks down), but here there is a singularity at the pole, so that this is a case ruled out. Let us, however, consider the example obtained by transforming the Euclidean line element:

$$x'^2 + y'^2 = T, \tag{35}$$

by putting

$$x = q_1, \quad y = q_2 + aq_1q_2,$$

so that the line element becomes

$$T = (1 + a^2q_2^2)q_1'^2 + 2aq_2(1 + aq_1)q_1'q_2' + (1 + aq_1)^2q_2'^2, \tag{36}$$

$$\sqrt{g} = \sqrt{|g_{ik}|} = 1 + aq_1,$$

$$T = p_1^2 - \frac{2aq_2}{1 + aq_1}\, p_1p_2 + \frac{1 + a^2q_2^2}{(1+aq_1)^2}\, p_2^2.$$

The difference between (29) and (30) in this case is:

$$\frac{h^2}{4\pi^2}\left[\frac{a}{1 + aq_1}\frac{\partial}{\partial q_1} - \frac{2a^2q_2}{(1 + aq_1)^2}\frac{\partial}{\partial q_2} - \frac{4a^2}{(1 + aq_1)^2}\right],$$

a non-zero operator, in general, except when $a = 0$, so that the restriction just mentioned does not help matters.

It is perhaps worth noting that the objections to assumption (2) no longer apply if we restrict T_1 and T_2 to have no common variables. In this case the result of applying (30) to each part

separately and adding is the same as that obtained by direct application to the whole form. This is what Podolsky[9] has done, although he states assumption (2) without any restriction.

8. Conclusions. The preceding critical comparison of the theories of the correspondence of operators and Hamiltonian functions leads us to the conclusion that the methods put forth have not the generality implied and claimed for them.

The fundamental correspondence (31), and (30) applied to obtain point transformations of it are all that are used in any applications known to the writer, and so long as attention is confined to these cases, the situation is clear.

In classical mechanics, we are so familiar with Lagrangian systems, invariant under point transformations, and Hamiltonian systems, invariant under contact transformations, that it is natural to attempt a similar generality of formulation at the outset for quantum mechanics. As we have indicated, two of the attempts to do this are not invariant under point transformations. Even the Schrödinger method gives no indication of the meaning of contact transformations in quantum mechanics. The variational approach to the problem is probably little more than the use of variational methods to effect transformations of point coördinates on the statement for the Cartesian case.

[9] B. Podolsky, *Phys. Review*, vol. 32, p. 812 (1928).

GEOMETRY OF A COMPLEX OF CURVES IN A RIEMANNIAN SPACE OF n DIMENSIONS

By C. L. E. Moore

1. Introduction. In a series of notes in 1887, 1888, 1889 Abbe Issaly [1] discussed what he called *pseudo surfaces* in a flat space of three dimensions. They are defined as follows:

$$dx = Pdu + Qdu',$$
$$dy = P'du + Q'du',$$
$$dy = P''du + Q''du',$$

where dx, dy, dz are *not* exact differentials. The P's and Q's are functions of u and u' only. The element of arc is then

$$ds^2 = Adu^2 + B'du'^2 + 2Bdudu'.$$

At a point P the directions defined lie in a plane. There is then a congruence of curves normal to these planar elements and the geometry then is that of the complex of curves normal to this special congruence. In case dx, dy, dz are exact differentials this reduces to the geometry of a surface in ordinary 3-space. Lines of curvature, geodesic lines and asymptotic lines are defined for the pseudo surface. The same author has written other papers on the same subject.[2]

In 1901 Dall'Aqua [3] discussed the general problem of the geometry of the complex of curves orthogonal to a general congruence in a Riemannian space of three dimensions. Here also lines of curvature, geodesic lines and asymptotic lines are defined. Issaly used the methods of Darboux while Dall'Aqua used the Absolute Calculus of Ricci. Curiously enough the

[1] Bulletin des Sciences mathematique.
[2] Nouvelles Annales de Mathematique, 1890, 1901.
[3] Sulla theoria delle congruenze di curve in una varieta qualunque a tre dimensioni. Annali di Matematica, vol. IV, series III.

latter derived the condition that two vectors be parallel, which agrees with the Levi-Civita definition. He bases his argument on a remark of Poincaré for which no reference is given. It seems to me, however, that his condition is only a necessary one, while he uses it as if it were both necessary and sufficient. If he had used his condition as the definition of parallelism, he certainly would have antedated Levi-Civita's work. Many other parts of the paper can be justified only on the basis of the parallelism of Levi-Civita.

In the present note I discuss the geometry of the complex of curves orthogonal to a general congruence in a Riemannian space of n dimensions. I define the covariant derivative with respect to the complex, parallelism in the complex, etc. Tensors similar to the Riemann-Christoffel tensor are found, and curvature of the complex thereby defined. I also introduce a second fundamental form which is quite analogous to the second fundamental form of a hypersurface. The generalization of the Gauss-Codazzi relations is also given. The latter, however, contain the second as well as the first derivatives.

2. The first fundamental form. Let the element of arc in the Riemannian space V_n be

$$ds^2 = a_{rs}dx^r dx^s \tag{1}$$

and let the congruence be $\lambda_{n|}$. Then the complex which we shall call the complex $\lambda_{n|}$ is defined by the Pfaffian equation

$$\lambda_{n|r}dx^r = 0. \tag{2}$$

Taking a set of $n-1$ congruences which form, with $\lambda_{n|}$, an orthogonal n-tuple, we have the following well known relations

$$\begin{aligned} \sum_1^n{}_i \lambda_{i|r}\lambda_{i|s} &= a_{rs} \\ \sum_1^n{}_i \lambda_{i|r}\lambda_{i|}{}^s &= \epsilon_r{}^s \\ \lambda_{i|r}\lambda_{j|}{}^r &= \epsilon_{ij}. \end{aligned} \tag{3}$$

Ricci's coefficients of relation are then defined by

$$\gamma_{ijk} = \lambda_{i|r\,s}\lambda_{j|}{}^{r}\lambda_{k|}{}^{s} \tag{4}$$

or

$$\lambda_{i|r\,s} = \sum_{1}^{n}{}_{jk}\, \gamma_{ijk}\lambda_{j|r}\lambda_{k|s}.$$

If now we subtract the square of (2) from (1) we get the linear element of the complex

$$ds'^{2} = a'_{r\,s}dx^{r}dx^{s} \tag{5}$$

where

$$a'_{r\,s} = a_{r\,s} - \lambda_{n|r}\lambda_{n|s} = \sum_{1}^{n-1}{}_{i}\, \lambda_{i|r}\lambda_{i|s}. \tag{6}$$

We do not assume that (2) is integrable and therefore we cannot, in general, eliminate one of the variables from the linear element (5). The tensor (6) is, however, singular as in the case of the first fundamental form for a hypersurface in V_n. If $\lambda_{n|}$ is contracted with $a'_{r\,s}$ the result is zero. That is, if any vector is contracted with $a'_{r\,s}$ the resulting vector has no component in the direction $\lambda_{n|}$. We will then say, as usual, that contracting a vector v^{r} with $a'_{r\,s}$ projects that vector on the complex and expresses the result in covariant form. It is at once evident that this result is accomplished in contravariant form by contracting with

$$\Delta_{r}{}^{s} = \sum_{1}^{n-1}{}_{i}\, \lambda_{i|r}\lambda_{i}{}^{s}. \tag{7}$$

The projection of any tensor can then be accomplished by contracting with the proper number of these tensors. We will adopt the notation

$$\Delta^{r\,s\,t}_{l\,m\,n} = \Delta_{l}{}^{r}\Delta_{m}{}^{s}\Delta_{n}{}^{t}. \tag{8}$$

Using the prime to denote the projected tensor we have

$$X'^{\,l\,t}_{r\,s} = \Delta^{\alpha\beta\,t}_{r\,s\,\gamma}X^{\gamma}_{\alpha\beta}.$$

Since any vector or tensor can be expressed in terms of the λ'^{s}

it is evident that contracting a vector or tensor which lies wholly in the complex, i.e. has no component containing $\lambda_{n|}$, with Δ does not change it.

3. The tensor $\lambda'_{n|rs}$. From equation (4) we have

$$\lambda_{n|rs} = \sum_{1}^{n}{}_{i,k} \gamma_{nik}\lambda_{i|r}\lambda_{k|s}.$$

Since

$$\gamma_{ijk} = -\gamma_{jik},$$

$$\lambda_{n|rs} = \sum_{1}^{n-1}{}_{i,j} \gamma_{nij}\lambda_{i|r}\lambda_{j|s} + \lambda_{n|s}\sum_{1}^{n-1}{}_{i} \gamma_{nin}\lambda_{i|r}$$

$$= \sum_{1}^{n-1}{}_{i,j} \gamma_{nij}\lambda_{i|r}\lambda_{j|s} - \lambda_{n|s}c_r,$$

where $c_r = -\sum_i \gamma_{nin}\lambda_{i|r}$ is the geodesic curvature[4] of the curve of the congruence $\lambda_{n|}$ which passes through the point considered. Hence we can write

$$\Omega_{rs} = \lambda'_{n|rs} = \sum_{1}^{n-1}{}_{i,j}\gamma_{nij}\lambda_{i|r}\lambda_{j|s} \tag{9}$$

The conditions that the congruence $\lambda_{n|}$ be normal are

$$\gamma_{nhk} = \gamma_{nkh} \quad h, k = 1, 2 \cdots n-1,$$

from which we see at once that the tensor Ω_{rs} is symmetric when and only when $\lambda_{n|}$ is a normal congruence.[5] We will see that this tensor is the one which gives us the second fundamental form for the complex $\lambda_{n|}$. We write the symmetric and anti-symmetric parts of $\lambda_{n|rs}$ as follows

$$\tfrac{1}{2}(\Omega_{rs} + \Omega_{sr}) = \tfrac{1}{2}\omega_{rs},$$
$$\tfrac{1}{2}(\Omega_{rs} - \Omega_{sr}) = \tfrac{1}{2}\omega'_{rs}.$$

4. The covariant derivative with respect to the complex. It is well known that the covariant derivative of a tensor lying in a hypersurface with respect to the hypersurface is the projection

[4] Ricci and Levi-Civita, Methodes de Calcul differentiel absolu, *Math. Ann.*, vol. 54, 1901, p. 154.

[5] Struik, *loc. cit.*, p. 55.

on the hypersurface of the covariant derivative with respect to the V_n.[6] We will then define the covariant derivative of a vector or tensor lying in the complex, as the projection of the derivative with respect to the V_n. We will find it more convenient to adopt the notation

$$\nabla_s v^t = v_s{}^t$$
$$\nabla_s' v^t = v_s'{}^t$$

for the derivative of a vector v (lying in the complex) with respect to V_n and the complex respectively. The same notation will be used for the derivatives of any tensor lying in the complex. If p is any scalar function of the λ_s' we have

$$\nabla_s' p = \Delta_s{}^t \nabla_t p. \tag{10}$$

and if $X_{rs}{}^t$ is any mixed tensor lying in the complex

$$\nabla_u' X_{rs}^t = \Delta_{rsi\gamma u}^{\alpha\beta t\ b} \nabla_\beta X_{\alpha\beta}^\gamma \tag{11}$$

Then

$$\begin{aligned} \nabla_t' a_{rs}' &= \Delta_{rsi}^{\alpha\beta\gamma} \nabla_\gamma a_{\alpha\beta}' = \Delta_{rsi}^{\alpha\beta\gamma} \nabla_\gamma (a_{\alpha\beta} - \lambda_{n|\alpha} \lambda_{n|\beta}) \\ &= -\Delta_{rsi}^{\alpha\beta\gamma} \nabla_\gamma (\lambda_{n|\alpha} \lambda_{n|\beta}) = 0. \end{aligned} \tag{12}$$

In a similar manner

$$\nabla_t' a'^{rs} = -\Delta_{\alpha\beta i}^{rs\gamma} \nabla_\gamma \lambda_n^\alpha | \lambda_n^\beta | = 0. \tag{12'}$$

The operator ∇' obeys the same laws as ∇ when operating on products. Then,

$$\nabla_t' \Delta_s{}^r = \nabla_t' a_{\alpha s}' a'^{\alpha r} = 0. \tag{12''}$$

Another formula we shall have use for is

$$\begin{aligned} \nabla_t \Delta_s{}^r &= \nabla_t (a_s{}^r - \lambda_{n|s} \lambda_n |^r) \\ &= -\lambda_{n|st} \lambda_n |^r - \lambda_{n|s} \lambda_n^s |_t. \end{aligned} \tag{13}$$

For the second derivative of a scalar function p we have

$$\begin{aligned} \nabla_s' \nabla_r' p &= \Delta_{sr}^\beta \nabla_\beta (\Delta_\gamma{}^\alpha \nabla_\alpha p) \\ &= \Delta_{sr}^\beta \Delta_\gamma{}^\alpha \nabla_\beta \nabla_\alpha p - \Delta_{sr}^\beta (\lambda_{n|\gamma\beta} \lambda_n^\alpha | + \lambda_{n|\gamma} \lambda_n^\alpha |_\beta) \nabla_\alpha p \\ &= \Delta_{sr}^{\beta\alpha} \nabla_\beta \nabla_\alpha p - \Delta_{sr}^{\beta\gamma} \lambda_{n|\gamma\beta} \lambda_n^\alpha | \Delta_\alpha p. \end{aligned}$$

[6] Struik, Mehr dimensionalen Differenzial geometrie, Berlin, Springer, 1922, p. 91.

Likewise

$$\nabla_r'\nabla_s'p = \Delta^{\beta\alpha}_{sr}\nabla_\alpha\nabla_\beta p - \Delta^{\beta\gamma}_{sr}\lambda_{n|\beta\gamma}\lambda^\alpha_{n|}\nabla_\alpha p.$$

Hence

$$\begin{aligned}\nabla_s'\nabla_r'p - \nabla_r'\nabla_s'p &= \nabla_s' \times \nabla_r'p \\ &= \Delta^{\beta\gamma}_{sr}(\lambda_{n|\beta\gamma} - \lambda_{n|\gamma\beta})\lambda^\alpha_{n|}\nabla_\alpha p \\ &= \omega'_{sr}\lambda^\alpha_{n|}\nabla_\alpha p.\end{aligned} \tag{14}$$

The right member of this equation vanishes when and only when $\lambda_{n|}$ is a normal congruence and in this case the problem reduces to that of the geometry of a hypersurface in V_n.

5. Making use of this last equation we can now get the second covariant derivative of any vector lying in the complex,

$$\begin{aligned}\nabla_s' \times \nabla_r'v^i &= \sum_1^{n-1} {}_i\, \nabla_s' \times \nabla_r'[\lambda_{i|\alpha}v^\alpha\lambda^i_{|}] \\ &= \sum_1^{n-1} {}_i\, (\nabla_s' \times \nabla_r'\lambda^i_{|})\lambda_{i|\alpha}v^\alpha + \sum_1^{n-1} {}_i(\nabla_s' \times \nabla_r'\lambda_{i|\alpha}v^\alpha)\lambda^i_{|} \\ &= \sum_1^{n-1} {}_i\, (\nabla_s' \times \nabla_r'\lambda^i_{|})\lambda_{i|\alpha}v^\alpha + \sum_1^{n-1} {}_i\, \omega'_{sr}\lambda^\beta_{n|}(\nabla_\beta\lambda_{i|\alpha}v^\alpha)\lambda^i_{|} \\ &= \sum_1^{n-1} {}_i\, (\nabla_s' \times \nabla_r'\lambda^i_{|})\lambda_{i|\alpha}v^\alpha + \sum_1^{n-1} {}_i\omega'_{sr}\lambda^\beta_{n|}(\nabla_\beta\lambda_{i|\alpha})\lambda^i_{|\alpha}v^\alpha \\ &\qquad + \omega'_{sr}\lambda^\beta_{n|}(\nabla_\beta v^\alpha)\lambda_{i|\alpha}\lambda^i_{|}\end{aligned}$$

Transposing we obtain the following identity

$$\begin{aligned}\nabla_s' \times \nabla_r'v^i - \sum_1^{n-1} {}_i\, \omega'_{sr}\lambda^\beta_{n|}(\nabla_\beta v^\alpha)\lambda_{i|\alpha}\lambda^i_{|} \\ = \sum_1^{n-1} {}_i\, [(\nabla_s' \times \nabla_r'\lambda^i_{|})\lambda_{i|\alpha} + \omega'_{sr}\lambda^\beta_{n|}(\nabla_\beta\lambda_{i|\alpha})\lambda^i_{|}]v^\alpha.\end{aligned}$$

The quantity $\sum \lambda_{i|\alpha}\lambda_{i|}{}^\iota = \Delta_\alpha{}^\iota$. We can obviously write the second term in the brackets on the right

$$\omega'_{sr}\lambda^\beta_{n|}(\nabla_\beta\lambda_{i|\rho})\Delta_\alpha{}^\rho\lambda^i_{|}$$

since contracting this with v^α gives the same result. This has the advantage that the term in the brackets now has no normal

component. We then write this equation in the form

$$(15)\quad \nabla_s' \times \nabla_r' v^i - \omega'_{sr}\lambda^\beta_{n|}\Delta_\alpha{}^i\nabla_\beta v^\alpha = \sum_1^{n-1} {}_i\,[(\nabla_s' \times \nabla_r'\lambda'_{i|})\lambda_{i|\alpha} + \omega'_{sr}\lambda^\beta_{x|}\Delta_\alpha{}^\rho(\nabla_\beta\lambda_{i|\rho})\lambda'_{i|}]v^\alpha$$

Likewise we obtain the similar formula

$$(16)\quad \nabla_s' \times \nabla_r' v_i - \omega'_{sr}\lambda^\beta_{n|}\Delta_i{}^\alpha\nabla_\beta v_\alpha = \sum_1^{n-1} {}_i\,(\nabla_s' \times \nabla_r'\lambda_{i|i})\lambda^\alpha_{i|} + \omega'_{sr}\lambda^\beta_{n|}\Delta_\rho{}^\alpha\lambda_{i|i}\nabla_\beta\lambda^\rho_{i|}]v_\alpha.$$

These formulas are the generalizations, for the complex, of Ricci's identities for V_n,

$$\nabla_s \times \nabla_r v_k = R_{rsk}{}^l v^\alpha, \quad \nabla_s \times \nabla_r v_\alpha = R^\alpha_{rsi} v_\alpha.$$

The expression on the left obviously does not depend on the choice of the congruences which form with $\lambda_{n|}$ an orthogonal n-tuple and consequently the bracket on the right does not depend on this choice.

6. Generalization of Riemann's tensor. The identities of Ricci given above can be expressed as follows:[7]

$$\nabla_s \times \nabla_r v^i = \sum_1^n {}_i\,(\nabla_s \times \nabla_r \lambda_{i|\alpha}\lambda^i_{i|}v^\alpha)$$

$$= \sum_1^n {}_i\,(\nabla_s \times \nabla_r \lambda_i{}^i)\lambda_{i|\alpha}v^\alpha$$

$$\nabla_s \times \nabla_r v_i = \sum_1^n {}_i\,(\nabla_s \times \nabla_r \lambda_{i|i})\lambda^\alpha_{i|}v_\alpha$$

and hence we have for the Riemann symbols of the second kind

$$R^i_{rsu} = \sum_1^n {}_i\,(\nabla_s \times \nabla_r\lambda^i_{i|})\lambda_{i|u},$$

$$R_{rsiu} = \sum_1^n {}_i\,(\nabla_s \times \nabla_r\lambda_{i|i})\lambda^u_{i|}.$$

[7] Struik, *loc. cit.*, p. 58.

By analogy with (15) and (16) we will define the symbols

$$(17)\qquad R'_{rstu} = \sum_1^{n-1}{}_i \left[(\nabla_s' \times \nabla_r'\lambda_{i|t})\lambda^u_{i|} - \omega'_{rs}\lambda^\beta_{n|}\lambda_{i|t}\Delta_\alpha{}^u\nabla_\beta\lambda^\alpha_{i|}\right]$$

$$(18)\qquad R'_{rstu} = a'_{\alpha u}R'^{\alpha}_{rst}$$
$$= \sum_1^{n-1}{}_i \left[(\nabla_s' \times \nabla_r'\lambda_{i|t})\lambda_{i|u} - \omega'_{rs}\lambda^\beta_{n|}\lambda_{i|t}\Delta_u{}^\alpha\nabla_\beta\lambda_{i|\alpha}\right]$$

In (18) the second term on the right can be written

$$\omega'_{rs}\lambda^\beta_{n|}\lambda_{i|t}\Delta_u{}^\alpha\lambda_{i|\alpha\beta} = \omega'_{rs}\sum_1^{n-1}{}_{i,\,j}\,\gamma_{ijn}\lambda_{i|t}\lambda_{j|u}$$
$$= \omega'_{rs}\pi_{tu}$$

where

$$\pi_{tu} = \sum_1^{n-1}{}_{i,\,j}\,\gamma_{ijn}\lambda_{i|t}\lambda_{j|u}$$

From the property of the symbols γ_{ijn} we see at once that the tensor π_{tu} is antisymmetric. An easy calculation, however, shows that this tensor is dependent upon the choice of the congruences $\lambda_{i|}$ which with $\lambda_{n|}$ form an orthogonal n-tuple. We can now write (18) in the form

$$(18')\qquad R'_{rstu} = \sum_1^{n-1}{}_i\,(\nabla_s' \times \nabla_r'\lambda_{i|t})\lambda_{i|u} + \omega'_{rs}\pi_{tu}.$$

8. Generalization of the Gauss-Codazzi relations. In the case of the hypersurface the Gauss formula expresses the relation between the Riemann symbols for the hypersurface and the projection on the hypersurface of the Riemann symbols for the V_n.[8] To obtain this relation for the complex we will first obtain the expansion for the second derivatives

$$\sum_1^{n-1}{}_i\,(\nabla_\mu'\nabla_\nu'\lambda_{i|r})\lambda_{i|s} = \sum_1^{n-1}{}_i\,\Delta^{\alpha\beta\gamma t}_{\mu\nu rs}\nabla_\alpha(\Delta^{\sigma\rho}_{\beta\gamma}\nabla_\sigma\lambda_{i|\rho})\lambda_{i|t}$$
$$= \sum_1^{n-1}{}_i\,\Delta^{\alpha\sigma\rho t}_{\mu\nu rs}(\nabla_\alpha\nabla_\sigma\lambda_{i|\rho})\lambda_{i|t} + \sum_1^{n-1}{}_i\,\Delta^{\alpha\beta\gamma}_{\mu\nu r}(\nabla_\alpha\Delta^{\sigma\rho}_{\beta\gamma})(\nabla_\sigma\lambda_{i|\rho})\lambda_{i|s}$$

[8] Struik, *loc. cit.*, p. 123.

$$= \sum_1^{n-1}{}_i \Delta^{\alpha\sigma\rho t}_{\mu\nu r s}(\nabla_\alpha \nabla_\sigma \lambda_{i|\rho})\lambda_{i|t} + \sum_1^{n-1}{}_i \Delta^{\alpha\beta\gamma}_{\mu\nu r}[(\nabla_\alpha \Delta_\beta{}^\sigma)^\rho \Delta_\gamma{}^\rho + (\nabla_\alpha \Delta_\gamma{}^\rho)\Delta_\beta{}^\sigma]\nabla_\sigma \lambda_{i|\rho}\lambda_{i|s}$$

$$= \sum_1^{n-1}{}_i \Delta^{\alpha\sigma\rho t}_{\mu\nu r s}(\nabla_\alpha \nabla_\sigma \lambda_{i|\rho})\lambda_{i|t} + \sum_1^{n-1}{}_i [\Delta^{\alpha\beta\rho}_{\mu\nu r}\nabla_\alpha \Delta_\beta{}^\sigma + \Delta^{\alpha\sigma\gamma}_{\mu\nu r}\nabla_\alpha \Delta_\gamma{}^\rho]\lambda_{i|s}\nabla_\sigma \lambda_{i|\rho}.$$

Expanding the terms of the last bracket separately we have

$$(19) \qquad \Delta^{\alpha\beta\rho}_{\mu\nu r}\nabla_\alpha \Delta_\beta{}^\sigma = -\Delta^{\alpha\beta\rho}_{\mu\nu r}\nabla_\alpha(\lambda_{n|\beta}\lambda^\sigma_{n|}) = -\Delta^{\alpha\beta\rho}_{\mu\nu r}(\nabla_\alpha \lambda_{n|\rho})\lambda^\sigma_{n|}$$
$$= -\Delta_r{}^\rho \Omega_{\nu\mu}\lambda^\sigma_{n|}.$$

$$(20) \qquad \Delta^{\alpha\sigma\gamma}_{\mu\nu r}\nabla_\alpha \Delta_\gamma{}^\rho = -\Delta^{\alpha\sigma\gamma}_{\mu\nu r}(\nabla_\alpha \lambda_{n|\gamma}\lambda^\rho_{n|}) = -\Delta^{\alpha\sigma\gamma}_{\mu\nu r}(\nabla_\alpha \lambda_{n|\gamma})\lambda^\rho_{n|}$$
$$= -\Omega_{r\mu}\Delta_\nu{}^\rho \lambda^\rho_{n|}.$$

Multiplying (19) by the term outside the brackets

$$-\Delta_r{}^\rho \Omega_{\nu\mu}\lambda^\sigma_{n|}\sum_1^{n-1}{}_i \lambda_{i|s}\nabla_\sigma \lambda_{i|\rho}$$
$$= -\Omega_{\nu\mu}\sum_1^{n-1}{}_{i,j}\sum_1^n{}_k \gamma_{ijk}\lambda_{j|r}\lambda_{k|\sigma}\lambda_{i|s}\lambda^\sigma_{n|}$$
$$= -\Omega_{\nu\mu}\pi_{sr} - \Omega_{r\mu}\Delta_\nu{}^\rho \lambda^\rho_{n|}\sum_1^{n-1}{}_{i|s}\nabla_\sigma \lambda_{i|\rho} = \Omega_{r\mu}\Omega_{s\nu}$$

Substituting these values we can then write the product

$$\sum_1^{n-1}{}_i (\nabla_\mu' \times \nabla_\nu' \lambda_{i|r})\lambda_{i|s}\sum_1^{n-1}{}_i \Delta^{\alpha\sigma\rho t}_{\mu\nu r s}R_{\alpha\sigma\rho t} + \Omega_{r\mu}\Omega_{s\nu}\Omega_{r\nu}\Omega_{s\mu} + \omega'_{r\mu}\pi_{rs}$$

and consequently we can write

$$(21) \qquad R'_{\mu\nu r s} = \Delta^{\alpha\sigma\rho t}_{\mu\nu r s}R_{\alpha\sigma\rho t} + \Omega_{r\mu}\Omega_{s\nu} - \Omega_{r\nu}\Omega_{s\mu}$$

which is the generalization of the Gauss relation. Using the identities which the symbols R satisfy we see from (21) that the symbols R' satisfy the two identities

$$(22) \qquad R'_{\mu\nu r s} = -R'_{\nu\mu r s}$$
$$R'_{\mu\nu r s} = -R'_{\mu\nu s r}.$$

Since Ω is not a symmetric tensor it seems obvious that the other identities satisfied by R are not satisfied by R'.

We can proceed in a similar manner to obtain the generalization of Codazzi's relations. If we project the Ricci identity

$$\nabla_t \times \nabla_s \lambda_{n|r} = R^m_{st\,r} \lambda_{n|\alpha}$$

on the complex we have

$$\Delta^{trs}_{\gamma\beta\alpha} R^m_{st\,r} \lambda_{n|\alpha} = \Delta^{trs}_{\gamma\beta\alpha} \nabla_t \times \nabla_s \lambda_{n|r}.$$

In order to obtain the expression for the right member in terms of Ω and $\lambda_{n|}$ consider the following expression

$$\begin{aligned}
\nabla_j' \Omega_{\alpha\beta} - \nabla_\beta' \Omega_{\alpha\gamma} &= \Delta^{trs}_{\gamma\alpha\beta} (\nabla_t \Omega_{rs} - \nabla_s \Omega_{rt}) \\
&= \Delta^{trs}_{j\alpha\beta} (\nabla_t \Delta^{pq}_{rr} \nabla_p \lambda_{n|q} - \nabla_s \Delta^{pq}_{tr} \nabla_p \lambda_{n|q}) \\
&= \Delta^{trs}_{\gamma\alpha\beta} \{ \nabla_t \times \nabla_s \lambda_{n|r} + (\nabla_t \Delta^{pq}_{rr} - \nabla_s \Delta^{pq}_{tr}) \nabla_p \lambda_{n|q} \} \\
&= \Delta^{trs}_{\gamma\alpha\beta} \{ R^p_{rst} \lambda_{n|p} + [(\nabla_t \Delta_s{}^p) \Delta_r{}^q + (\nabla_t \Delta_r{}^q) \Delta_s{}^p - (\nabla_s \Delta_s{}^p) \Delta_r{}^q \\
&\qquad - (\nabla_s \Delta_r{}^q) \Delta_t{}^p] \nabla_p \lambda_{n|q} \} \\
&= \Delta^{trs}_{\gamma\alpha\beta} \{ R^p_{rst} \lambda_{n|p} + \omega'_{st} \lambda^p_{n|} \Delta_r{}^q \nabla_p \lambda_{n|q} \} \\
&= \Delta^{trs}_{\gamma\alpha\beta} R^p_{rst} \lambda_{n|p} + \omega'_{\beta\gamma} \lambda_n{}^p \Delta_r{}^q \nabla_p \lambda_{n|q}
\end{aligned}$$

Hence the relation sought is

$$\nabla_\gamma' \Omega_{\alpha\beta} - \nabla_\beta' \Omega_{\alpha\gamma} - \omega'_{\beta\gamma} \lambda^m_{p|} \Delta_r{}^q \nabla_p \lambda_{n|q} = \Delta^{trs}_{\gamma\alpha\beta} R^p_{rst} \lambda_{n|p} \tag{23}$$

which is the generalization of Codazzi's relations. For hypersurfaces this relation involves only the $\lambda_{n|}$ and Ω but here we see that the derivatives of $\lambda_{n|}$ are also involved.

8. Curvature of the Complex. We will define the curvature of the complex in any orientation as

$$K_{ij} = R'_{rstu} (\mu_i{}^r \times \mu_j{}^s)(\mu_i{}^t \times \mu_j{}^u)$$

where μ_i and μ_j are two mutually orthogonal unit vectors of the complex which define the orientation at the point considered. If we choose the congruences λ_i so that μ_i is tangent to λ_i and μ_j is tangent to λ_j, then we can write

$$K_{ij} = R'_{rstu} (\lambda^r_{i|} \times \lambda^s_{j|})(\lambda^t_{i|} \times \lambda^u_{j|}) \tag{24}$$

The quantity $\lambda^r_{i|} \times \lambda^s_{j|}$ is the unit area in the plane of the

orientation. Now using identities (21) which R' satisfies

$$K_{ij} = R'_{rstu}(\lambda^r_{i|}\lambda^s_{j|}\lambda^t_{i|}\lambda^u_{j|} - \lambda^r_{i|}\lambda^s_{j|}\lambda^t_{j|}\lambda^u_{i|} - \lambda^s_{j|}\lambda^t_{i|}\lambda^u_{j|} + \lambda^r_{j|}\lambda^s_{i|}\lambda^t_{j|}\lambda^u_{i|}) = 4R'_{rstu}\lambda^r_{i|}\lambda^s_{j|}\lambda^t_{i|}\lambda^u_{j|}.$$

We saw that R' did not depend on the choice of the congruence $\lambda_{i|}$ consequently K_{ij} is really an invariant of the complex. This might be considered as the generalization of the Riemannian curvature.

Another invariant is

$$\begin{aligned} K &= a'^{rs}a'^{tu}R'_{rstu} \\ &= \sum_1^{n-1}{}_{i,j}\,\lambda^r_{i|}\lambda^s_{i|}\lambda^t_{j|}\lambda^u_{j|}R'_{rstu} \\ &= \tfrac{1}{4}\sum_1^{n-1}{}_{i,j}\,(\lambda^r_{i|}\times\lambda^s_{j|})(\lambda^t_{i|}\times\lambda^u_{j|})R'_{rstu} \\ &= \sum_1^{n-1}{}_{i,j}\,K_{ij}. \end{aligned} \tag{26}$$

Hence K is the sum of all the curvatures of the orientations formed by the $n-1$ curves forming the orthogonal n-tuple. This is the mean curvature for the point P. A third invariant is

$$\begin{aligned} K_i &= \sum_1^{n-1}{}_j\,K_{ij} = 4\textstyle\sum_j R'_{istu}\,\lambda^r_{i|}\lambda^s_{j|}\lambda^t_{i|}\lambda^u_{j|} \\ &= R'_{rstu}\lambda^r_{i|}\lambda^t_{i|}a'^{su} = R'_{ri}\lambda^r_{i|}\lambda^t_{i|}. \end{aligned} \tag{27}$$

This invariant depends on the congruence $\lambda_{i|}$ and not on the other $n-2$ congruences of the n-tuple. It is the sum of all the curvatures of all the orientations formed by $\lambda_{i|}$ and the other $n-2$ directions tangent to the n-tuple. This is the mean curvature of the complex with respect to a given direction through the point considered.

The tensor R_{rt}' defined by (27) can now be used to determine $n-1$ mutually perpendicular directions through each point of the complex which are analogous to the principal directions of Ricci.

9. Parallelism. In V_n the covariant differential is

$$\delta v_r = v_{r,s}dx^s$$

and if v is a vector of the complex we will define

$$\begin{aligned}\delta' v_r &= \nabla_s' v_r dx^s \\ &= \Delta_{rs}^{\alpha\beta}\Delta_\beta v_\alpha fx^s \\ &= \Delta^\alpha{}_r \delta v_\alpha\end{aligned}$$

which shows that the covariant differential with respect to the complex is the projection on the complex of the differential with respect to V_n. Then when a vector of the complex is moved by parallelism in direction dx of the complex the increment added to v_r is $\delta' v_r$ and similarly the increment added to v^r is $\delta' v^r$. Hence *if a vector of the complex is moved by parallelism with respect to V_n, the projection is parallel with respect to the complex.* Also two vectors are parallel with respect to a curve C of the complex if they are parallel with respect to any surface tangent to the complex along C (perpendicular to the curves $\lambda_{n|}$ which intersect C).

If a vector d_1x^μ is moved by parallelism along d_2x^μ and then d_2x^μ is moved by parallelism along d_1x^μ we have

$$\delta_1' d_2x^\mu - \delta_2' d_1x^\mu = \Delta_\mu{}^\nu(\delta_1 d_2x^\mu - \delta_2 d_1x^\mu) = 0 \tag{28}$$

Hence we have a closed figure the opposite sides of which are parallel. To show that the opposite sides are equal, consider two vectors of the complex u_r, v^r and compare the product $v_r v^r$ before the displacement with the product after the displacement. From (28) if p is a scalar function

$$\delta' p = \Delta^s{}_r \nabla_r p dx^s = \nabla_r p dx^r = d'p$$

where dx^r lies in the complex. Also we can write

$$\begin{aligned}\delta' u_r &= \Delta_{ri}^{\alpha\iota}\nabla_\alpha u_\iota dx^s = \Delta_r{}^\iota \nabla_\alpha u_\iota dx^\alpha \\ &= \Delta_r{}^\iota\left(du_\iota - \left\{\begin{matrix}\iota\alpha \\ \beta\end{matrix}\right\} u^\beta dx^\alpha\right)\end{aligned} \tag{29}$$

Likewise

$$\delta' v^r = \Delta_t{}^r \left(dv^t + \left\{ \begin{matrix} \alpha\beta \\ t \end{matrix} \right\} v^\beta dx^\alpha \right) \tag{30}$$

If the vector fields u, v are parallel then

$$\delta' u_r = \delta' v^r = 0$$

and hence the increments are

$$d' u_r = \Delta_r{}^t \left\{ \begin{matrix} t\alpha \\ \beta \end{matrix} \right\} u_\beta dx^\alpha, \quad d' v^r = - \Delta_t{}^r \left\{ \begin{matrix} \alpha\beta \\ t \end{matrix} \right\} v^\beta dx^\alpha$$

Then after displacement

$$\begin{aligned} (u_r + d' u_r)(v^r + d' v^r) &= u_r v^r + v^r d' u_r + u_r d' v^r \\ &= u_r v^r + \Delta_r{}^t \left\{ \begin{matrix} t\alpha \\ \beta \end{matrix} \right\} u_\beta v^r dx^\alpha - \Delta_t{}^r \left\{ \begin{matrix} \alpha\beta \\ t \end{matrix} \right\} u_r v^\beta dx^\alpha \\ &= u_r v^r + \left\{ \begin{matrix} t\alpha \\ \beta \end{matrix} \right\} u_\beta v^t dx^\alpha - \left\{ \begin{matrix} \alpha\beta \\ t \end{matrix} \right\} u_t v^\beta dx^\alpha \\ &= u_r v^r. \end{aligned}$$

If u, v are the same vector the product $u^r u_r$ is the length and if u, v are unit vectors then $u_r v^r$ is the cosine of the angle between them. Hence if a vector is displaced by parallelism the length is unchanged and if two vectors at a point are displaced by parallelism the angle between them is left unchanged. Then the infinitesimal parallelogram above has its opposite sides equal.

Now move the vector v^r around the infinitesimal parallelogram determined by the points

$$x, \quad x + d_1 x, \quad x + d_1 x + d_2 x, \quad x + d_2 x.$$

We accomplish this by first moving v^r from x by parallelism along the first two sides and along the other two sides from x in the opposite direction along the other two sides and then taking the difference between the second and fourth positions.

(a) $v^r + \delta_1' v^r$

(b) $v^r + \delta_2' v^r + \delta_1' v^r + \delta_2' \delta_1' v^r$

(c) $\qquad v^r - \delta_2' v^r$

(d) $\qquad v^r + \delta_1' v^r + \delta_2' v^r + .$

Hence the total change in v^r due to this displacement is

$$\begin{aligned}\Delta v^r &= \delta_2'\delta_1' v^r - \delta_1'\delta_2' v^r = (\delta_2'\delta_1' - \delta_1'\delta_2')v^r \\ &= \nabla_i'\nabla_s' v^r d_1x^s d_2x^t - \nabla_i'\nabla_s' v^r d_2x^s d_1x^t \\ &= \nabla_i'\nabla_s'(d_1x^s d_2x^t - d_2x^s d_1x^t) \\ &= \nabla_i'\nabla_s' d_1x^s \times d_2x^t.\end{aligned}$$

Since $d_1x^s \times d_2x^t$ is antisymmetric we can write this

$$\Delta v^r = (\nabla_i' \times \nabla_s' v^r) d_1x^s \times d_2x^t$$

From (24) we see that the change in a vector by this displacement is not connected with the curvature as in the case of the V_n itself.

10. Geodesics. *The geodesic curvature* of a curve of the complex is defined by the relation

$$(31) \qquad c_r = \frac{\delta'\lambda_r}{ds} = (\nabla_s'\lambda_r)\lambda^s \quad \text{or} \quad c^r = \frac{\delta'\lambda^r}{ds} = \nabla_s'\lambda^r\lambda^s$$

where λ is a unit vector. A geodesic is then a curve of zero geodesic curvature. If we take the congruence λ_1 to coincide with λ we have

$$(32) \qquad \begin{aligned} c_r &= \Delta^{\alpha\beta}_{sr}(\nabla_\alpha \lambda_{i|\beta})\lambda^s_{i|} = \Delta^{\beta}_{r}(\nabla_\alpha \lambda_{i|\beta})\lambda^\alpha_{i|} \\ &= \Delta_r^{\ \beta} \sum_{j,k}^{n-1} \gamma_{ijk}\lambda_{j|\beta}\lambda_{k|\alpha}\lambda^\alpha_{i|} \\ &= \sum_{j}^{n-1} \gamma_{iji}\lambda_{j|r}.\end{aligned}$$

Hence if the congruence $\lambda_{i|}$ is a geodesic congruence then

$$(33) \qquad \gamma_{iji} = -\gamma_{jii} = 0, \quad j = 1, 2, \cdots, n-1.$$

It is not difficult to show that curves satisfying (32) are curves of shortest length.[9] It is well known that if the congruence $\lambda_{i|}$

[9] Dall'Aqua, *loc. cit.*

is a geodesic in V_n

$$\gamma_{ii|j} = 0, \quad j = 1, 2, \cdots, n.$$

Hence we have the following theorem

If a geodesic of V_n lies in the complex it is also a geodesic of the complex. If vectors of the complex are parallel in V_n with respect to a curve C of the complex they are also parallel in the complex with respect to C.

11. Normal curvature. The curvature in the V_n of a curve $\lambda_{i|}$ of the complex is

$$C_r = \lambda_{i|r\alpha}\lambda^r_{i|}. \tag{34}$$

The projection of this on the normal, that is along $\lambda_{n|}$, is

$$\lambda_{i|r\alpha}\lambda^\alpha_{i|}\lambda^r_{n|} = \gamma_{ini} = -\gamma_{nii} = -\lambda_{n|r\alpha}\lambda^r_{i|}\lambda^\alpha_{i|}.$$

Since it does not depend on any thing but $\lambda_{n|r\alpha}$, $\lambda^r_{i|}$, $\lambda^\alpha_{i|}$ we have the theorem: *The normal component of the curvature of all curves of the complex having the same tangent line is the same.* We call this the *normal curvature* of the complex in this direction.

Writing (34) in the form

$$\begin{aligned} C_r = \lambda_{i|r\alpha}\lambda^\alpha_{i|} &= \sum_{1}^{n}{}_{j,k}\,\gamma_{ijk}\lambda_{j|r}\lambda_{k|\alpha}\lambda^\alpha_{i|} \\ &= -\sum_{1}^{n}{}_{j}\,\gamma_{jii}\lambda_{j|r} = -\sum_{1}^{n-1}{}_{j}\,\gamma_{jii}\lambda_{i|r} - \lambda_{n|r}\gamma_{nii} \end{aligned}$$

Hence: *The curvature of a curves of the complex with respect to the V_n is equal to the sum of the geodesic and the normal curvatures.* From this it is seen that the curvature of a geodesic of the complex with respect to the V_n is normal to the complex and hence geodesics of the complex have the properties of geodesics on a hypersurface of V_n.

The *mean normal curvature* of the complex at a given point can be defined by

$$\begin{aligned} H = \sum_{1}^{n-1}{}_{i}\,\gamma_{nii} &= \sum_{1}^{n-1}{}_{i}\,\lambda_{n|r\alpha}\lambda^r_{i|}\lambda^\alpha_{i|} \\ &= \lambda_{n|r\alpha}\sum_{1}^{n-1}{}_{i}\,\lambda^r_{i|}\lambda^\alpha_{i|} = a^{r\alpha}\lambda_{n|r\alpha} \end{aligned} \tag{35}$$

Hence H is independent of the choice of the $n-1$ mutually orthogonal directions through the given point.

12. The second fundamental tensor of the complex. The derivative of $\lambda_{n|}$ with respect to V_n can now be written

$$\begin{aligned}\lambda_{n|r\sigma} &= \sum_{1}^{n}{}_{j,\,k}\,\gamma_{njk}\lambda_{j|r}\lambda_{k|\sigma} \\ &= \sum_{n}^{n-1}{}_{j,\,k}\,\gamma_{njk}\lambda_{j|r}\lambda_{k|\sigma} + \lambda_{n|\sigma}\sum_{1}^{n-1}{}_{j}\,\gamma_{njn}\lambda_{j|r} \\ &= \Omega_{r\sigma} + \lambda_{n|\sigma}\epsilon_{r}.\end{aligned} \tag{36}$$

where

$$\Omega_{r\sigma} = \sum_{1}^{n-1}{}_{j,\,k}\,\gamma_{njk}\lambda_{j|r}\lambda_{k|\sigma} \tag{37}$$

and is a tensor lying in the complex.[10] We will call this the *second fundamental tensor* of the complex. From (36) it is evident that $\Omega_{r\sigma}$ is not symmetric. It can also be written

$$\Omega_{r\sigma} = \Delta^{\alpha\beta}_{r\sigma}\lambda_{n|\alpha\beta}$$

and as the projection of a tensor does not depend upon the congruences in terms of which the complex is expressed we conclude that $\Omega_{r\sigma}$ does not depend upon the $n-1$ mutually orthogonal congruences λ_{i}. This can however be demonstrated by computing the same form for a new choice of congruence $\lambda_{i|}'$. It is found that the new form is identically equal to (37). The total set of Ricci's coefficients of rotation γ_{ijk} which have at least one subscript equal to n can be divided into three groups

$$\gamma_{nkj}, \quad \gamma_{njn}, \quad \gamma_{ijn}.$$

The first set determines the second fundamental tensor, the second determines the curvature of the curves $\lambda_{n|}$ and the third determined the tensor π_{ij} and from the expression for π_{ij} in terms of a new set of congruence $\lambda_{i|}'$ it is found to depend upon the choice of congruences $\lambda_{i|}$. It is seen however that π_{ij} cannot be made to vanish.

[10] Struik, *loc. cit.*, p. 56.

The normal curvature can now be written in the form

$$\rho_i = -\Omega_{rs}\lambda_{i|}{}^r\lambda_{i|}{}^s$$

and hence

$$C_r = c_r - \lambda_{i|}{}^r\lambda_{i|}{}^s\Omega_{rs}.$$

We can now decompose Ω into its symmetric and its anti-symmetric parts,

$$\begin{aligned}\Omega_{rs} &= \tfrac{1}{2}(\Omega_{rs} + \Omega_{sr}) + \tfrac{1}{2}(\Omega_{rs} - \Omega_{sr}) \\ &= \tfrac{1}{2}\omega_{rs} + \tfrac{1}{2}\omega_{rs}'.\end{aligned}$$

We then have a symmetric tensor

$$\omega_{rs} = \Omega_{rs} + \Omega_{sr}$$

and in terms of this can define principal directions in the usual way. That is, these congruences are determined as the solutions of the equations

$$(\omega_{rs} + \rho a_{rs}')\lambda^r = 0.$$

Since ω_{rs} and a_{rs}' have no component containing λ_n one solution will be $\lambda^r = \lambda_{n|}{}^r$ and there will be $n-1$ solutions which have no component along $\lambda_{|n}$. To show this, consider a point P and let coordinates be chosen so that $\lambda_{n|}$ is in the direction of the parameter curve X_n and the other parameter curves perpendicular to $\lambda_{n|}$. Then in the above equation r will take values $1, 2, \cdots, n-1$. This is then the system of equations which determine the principal directions as usually given. The invariants ρ_h determined by

$$|\omega_{rs} + \rho_h a_{rs}'| = 0$$

are the principal curvatures, for as usual we can write

$$(\omega_{rs} - \rho_h a_{rs}')\lambda_{h|}{}^r\lambda_{h|}{}^s = 0$$

and since

$$a_{rs}'\lambda_{h|}{}^r\lambda_{h|}{}^s = 1$$

we have

$$-\omega_{rs}\lambda_{h|}{}^r\lambda_{h|}{}^s = \rho_h.$$

This is the same as equation

$$\rho_k = \Omega_{rs}\lambda_{i|}{}^r\lambda_{i|}{}^s$$

because in the product $\Omega_{rs}\lambda_{i|}{}^r\lambda_{i|}{}^s$ the anti-symmetric part vanishes. It is evident that the values of ρ_k are the maximum and minimum values of the normal curvature.[11]

The congruences $\lambda_{k|}$, $\lambda_{k|}$ are orthogonal, and if we use these to form the orthogonal n-tuple

$$\omega_{rs}\lambda_{k|}{}^r\lambda_{k|}{}^s = 0$$

which requires that

$$\gamma_{nkk} + \gamma_{nkk} = 0, \quad h, k = 1, 2, \cdots, n-1.$$

and consequently this is Ricci's system cannonical with the congruence $\lambda_{n|}$. They envelope a set of mutually perpendicular curves of the complex. We define these as the *lines of curvature* of the complex. These lines do not have the property that normals to the complex at near by points of these curves intersect, so Dall'Aqua defines his lines of curvature to have this property. Consequently they are not mutually perpendicular. These latter directions can be defined according to my recent note [12] as the directions dx^r of the complex such that when λ_n is moved by parallelism, in the V_n, along dx^r the new position will lie in the planar element defined by $\lambda_{n|}$ and dx^r. There are $n-1$ of these directions and they are mutually perpendicular when and only when the congruence $\lambda_{n|}$ is normal and then they coincide with the canonical directions.

13. Asymptotic lines and conjugate directions. Asymptotic lines are defined as lines whose principal normal is tangent to the complex or curves whose normal curvature is zero. Hence the equation is

$$\Omega_{rs}\lambda_{i|}{}^r\lambda_{i|}{}^s = 0. \tag{37}$$

From (37) it is seen that the necessary condition that a con-

[11] Eisenhart, Riemannian Geometry, p. 108.

[12] C. L. E. Moore, "Note on Congruences in Riemannian Space," JOURNAL OF MATHEMATICS AND PHYSICS, vol. VII, p. 189.

gruence $\lambda_{i\hat{k}}$ be asymptotic is given by

$$\gamma_{jii} = 0.$$

This shows that: *If a geodesic of V_n lies in the complex it is an asymptotic line of the complex and conversely. If a line of the complex is a geodesic of V_n it must be an asymptotic line which is also a geodesic.*

The direction μ is said to be conjugate to $\lambda_{1|}$ if

$$\Omega_{rs}\lambda_{i|}{}^{r}\mu^{s} = 0 \tag{38}$$

but since Ω_{rs} is not symmetric λ_i will not, in general, be conjugate to μ. Now the tangent plane of the complex at a point is characterized by

$$\lambda_{n|r}\lambda_{i|}{}^{r} = 0.$$

The hyperplane passing through P, parallel to the hyperplane at $P + dP$ (where dP is in the direction μ) is characterized by the equation

$$\lambda_{n|rs}\lambda_{i|}{}^{r}\mu^{s} = 0$$

and since $\lambda_{i|}$ and μ are contained in the complex we can replace $\lambda_{n|rs}$ by Ω_{rs} and this equation coincides with (38). Hence: *The directions conjugate to a direction μ lie in the intersection of the tangent hyperplane and the parallel to the nearby tangent hyperplane in the direction μ.*

The relation between λ_1 and μ is symmetric when and only when $\lambda_{n|}$ is a normal congruence. Asymptotic directions are then self-conjugate directions.

SETTLEMENT OF BUILDINGS DUE TO PROGRESSIVE CONSOLIDATION OF INDIVIDUAL STRATA

By Charles Terzaghi

The problem which we attempt to solve is graphically represented in Fig. 1*a*. A feebly permeable, very compressible stratum

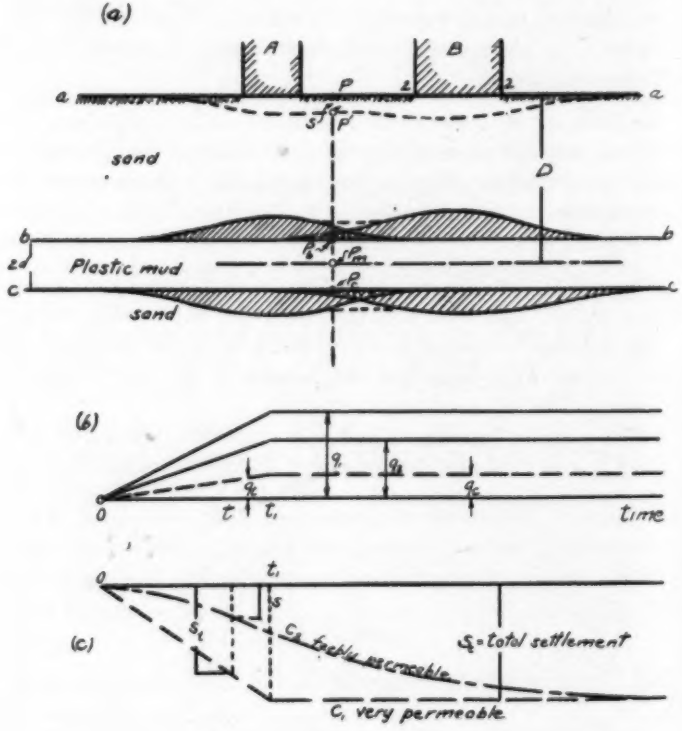


FIG. 1.

bbcc with a thickness $2d$ is located at a depth D below the surface *aa* of the ground. On the ground surface we construct buildings *A*, *B*, etc. Construction is started simultaneously at a time zero. As time goes on, the load acting on the areas 11, 22, etc. is supposed to increase at a uniform rate, until at a time t_1 after construction was started, the load assumed the values q_1, q_2, etc. per unit of area. After the time t_1 the load is supposed to remain constant. The diagram Fig. 1*b* which represents the time rate of load application (abscissae = time, ordinates = load acting on the surface of the ground) shall be called the "load progress diagram."

The investigation will be based on the following assumptions: (*a*) The layers of sand located above and below the compressible layer *bbcc* are very permeable and perfectly incompressible. (*b*) The compressible layer is homogeneous throughout and its thickness is everywhere the same. (*c*) The voids of the compressible layer are completely filled with water. (*d*) Before construction of the buildings *A*, *B*, etc. is started, the ground water is perfectly stationary, which in turn involves assuming that the water content of the compressible material everywhere corresponds to the pressure under which it stands. (*e*) If the pressure which acted on the compressible layer previous to construction (pressure due to the weight of the overburden) is increased by a certain amount, the corresponding compression gradually approaches a value which is equal to the increase in pressure times a constant. (*f*) The coefficient of permeability of the compressible layer is, within the range of the variation of the pressure, independent of the pressure. (*g*) The flow of water through the compressible layer strictly follows Darcy's law.

If assumption (*a*) does not strictly hold true, the subsidence of the structures will be equal to the subsidence caused by the compression of the layer *bbcc* plus the subsidence due to compression of the layers of sand. Since we are merely interested in the first item, the assumption (*a*) is justified. The importance of the error due to assumption (*b*) depends on the degree of uni-

formity of the natural soil deposit. Assumption (*c*) is almost invariable correct. Assumption (*d*) is almost correct, unless the layer *bbcc* is of very recent origin. Assumptions (*e*) and (*f*) are sufficiently accurate, provided the additional pressure exerted by the weight of the structures does not exceed the pressure exerted by the weight of the overburden *aabb*. As a matter of fact, the errors due to assumptions (*d*) and (*e*) partially compensate each other.[5] Assumption (*g*) is almost correct.

The following symbols will be used:

p_0 = the vertical pressure which acts at the point P_m before construction was started (pressure due to the weight of the overburden).

q_1, q_2, etc. load exerted by the buildings *A*, *B*, etc. per unit of area of the ground surface.

t_1 the time between start and finish of construction.

q_t the vertical pressure produced by the weight of the buildings *A*, *B*, etc. at the point P_m of the compressible layer at a time $t < t_1$ after construction was started.

q_e the vertical pressure produced by the weight of the buildings *A*, *B*, etc. at the point P_m of the compressible layer after construction was finished ($t > t_1$).

D average distance between the ground surface and the compressible layer.

$2d_1$ thickness of the compressible stratum.

k_1 coefficient of permeability of the compressible stratum (rate of percolation at a hydraulic gradient equal to unity).

a coefficient of compressibility of the compressible stratum, that means loss of water per unit of volume of dry material caused by an increase of the pressure q by unity.

s subsidence of the surface of the ground at a point *P*. This subsidence is assumed to be identical with the compression of the compressible layer beneath the point *P*.

s_t compression of the compressible layer at the point P_m (located beneath point *P*) at a time t after construction was started.

s_e compression of the compressible layer at the point P_m at a time $t = \infty$.

Since the subsidence $s = PP'$ of the ground surface at a point P is supposed to be equal to the compression of the compressible layer at the point P_m located beneath the point P, we have reduced our problem to the following question: What will be the compression of the layer *bbcc* beneath the point P, at an arbitrary time t after the application of the surface load was started?

To answer this question requires two independent operations. We must first of all find out the intensity of the vertical pressure q which acts on the compressible layer beneath the point P and then we have to find out the relation which exists between the pressure q, the compression s and the time t.

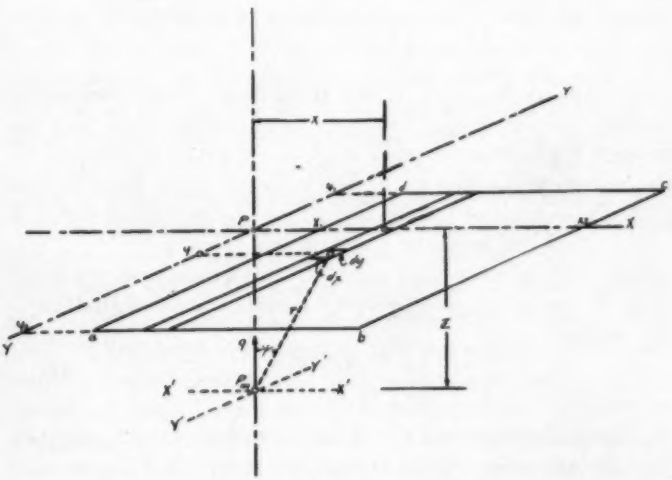


FIG. 2.

The first part of the investigation can easily be performed by means of the equations of Boussinesq. The area *abcd*, Fig. 2, represents one of the sites occupied by the buildings in Fig. 1*a*.

We divide this area into elements $dx\,dy$ and select the point P as origin of our coordinate system. The pressures which are produced by the load $q_1\,dx\,dy$ at the points P_b and P_c (Fig. 1*a*) are not identical, inasmuch as the pressure q will slightly increase from the top surface of the compressible layer towards the bottom surface. However, since the difference between these two pressures is not appreciable we can disregard it and assume that they are identical with the pressure acting at the point P_m halfway between P_b and P_c. The state of stress produced by the load $q_1\,dx\,dy$ is determined by the equations of Boussinesq.[1] By means of these equations one can compute for the point P the three principal stresses. Föppl [2] has shown that two of these principal stresses are so small that they can be neglected, while the third one, with the intensity

$$\frac{3}{2}\frac{q\,dx\,dy}{\pi}\frac{\cos\psi}{r^2}$$

passes through the point of load application. The experimental investigations of Kögler [3] seem to have demonstrated that the pressure distribution in sands is fairly accurately determined by the preceding formula for a depth greater than about 6 feet. Hence, if we assume t to be considerably greater than 6 feet, our computation can safely be based on this formula.

The principal stress determined by the formula produces at the point P_m on a horizontal element a vertical normal stress with the intensity

$$\frac{3}{2}\frac{q_1\,dx\,dy}{\pi}\frac{\cos^3\psi}{r^2}.$$

By integrating between the limits y_1 and y_2 (Fig. 2), we learn that the strip with a width dx produces at the point P_m a vertical normal stress with the intensity

$$dq = \frac{q_1 z^3}{2\pi(x^2+z^2)^4}\left[\frac{y_1(3x^2+2y_1^2+z^2)}{(x^2+y_1^2+z^2)^{\frac{1}{2}}} - \frac{y_2(3x^2+2y_2^2+z^2)}{(x^2+y_2^2+z^2)^{\frac{1}{2}}}\right]dx.$$

The second integration should preferably be performed by means

of a graphical procedure whereby the width Δx of the finite elements should be smaller than one-third of the depth z.

Thus the first part of our investigation has led to the knowledge of the average vertical pressure q which acts beneath the point P on our compressible layer. To make the step from the pressure to the compression, we must start with analysing the effect of the application of a vertical pressure q on the state of stress which exists in the compressible layer. From previous experimental investigations we know the following fact: If we raise the pressure p_0 which acts on a laterally confined mass of water-soaked soil by an amount q, the compression produced by q is approximately equal to the thickness $2d$ of the layer times the pressure q times a constant, provided q is smaller than p_0.[4] However, we also know that the compression by no means occurs at once. It proceeds at a decreasing rate and the ultimate value is not reached, except after a considerable period of time. An investigation [5] of this phenomenon has led to the conclusion that the lag is essentially due to the resistance against the escape of the excess water contained in the voids of the compressible material. If the compressible layer is very permeable, compression takes place almost at once and the settlement of point P would occur according to the broken line C_1, Fig. 1*c*. Since the surface load is applied at a uniform rate (Fig. 1*b*) the settlement would also increase at a uniform rate from time 0 until time t_1 and from time t_1 to $t = \infty$ they would remain constant. On the other hand, if the compressible layer is very feebly permeable, the settlement will be retarded, and instead of the settlement curve C_1 we would obtain the curve C_2 (Fig. 1*c*). The lower the permeability, the flatter this curve should be.

In order to determine the influence of permeability on the speed of compression and on the corresponding rate of settlement, we have to start from the following fundamental fact: Since the voids of our compressible stratum are supposed to be filled with water, the process of compression is obviously associated with the flow of water from the interior of the compressible

stratum into the adjoining permeable and incompressible medium. Water cannot flow through the voids of a soil unless there exists within the water a hydraulic gradient, that means a drop in hydrostatic pressure in the direction of the flow. Since the speed with which the water drains out of the stratum merely depends on the thickness of the stratum, on its permeability and on the hydraulic gradient, we are in a position to utilize the aforementioned hydraulic fact for computing the rate of compression of the compressible layer.

Before construction was started, the compressible layer was merely under the influence of the weight of the superimposed strata. This weight was called p_0 per unit of area of the horizontal plane which passes through the point p_m (Fig. 1, 2). Since p_0 is supposed to have acted on the layer for a very long time, the water content of the soil exactly corresponds to the pressure p_0 (our assumption *d*). Since no flow of water takes place, within the ground water, the hydrostatic excess pressure is everywhere equal to zero. Hence, the total weight p_0 of the overburden must be carried by the solid constituents of the compressible layer.

At a time t after the construction of the buildings A and B (Fig. 1*a*) was started, the pressure which acts at the point P_m on the compressible layer is equal to $p_0 + q_t$ (Fig. 1*b*). If the compressible layer was very permeable, the pressure $p_0 + q_t$ would entirely be carried by the solid constituents of this layer and the settlement caused by the application of q_t would be equal to s_t (Fig. 1*c*). However, due to the low permeability of the layer, the water content cannot drop down to the value which corresponds to the pressure $p_0 + q_t$ as rapidly as q_t was applied. Therefore the water in the central part of the layer comes under a pressure w. According to the laws of mechanics, the pressure $p_0 + q_t$ must in every part of the layer be equal to the sum of the pressure which acts on the solid constituents of the layer plus the pressure w which acts in the water, or, according to Fig. 3*a*

$$p_0 + q_t = p_0 + q + w,$$

or

$$q_t = q + w.$$

It is obvious that the values of w will be greatest near the center of the layer, because near the center the distance through which the excess water has to travel is a maximum. On the other hand, at the boundaries, w is equal to zero, because there the water can escape almost at once. From previous investigations we know that consolidation proceeds from the boundaries towards the center as shown in Fig. 3*a*. Starting at the boundary

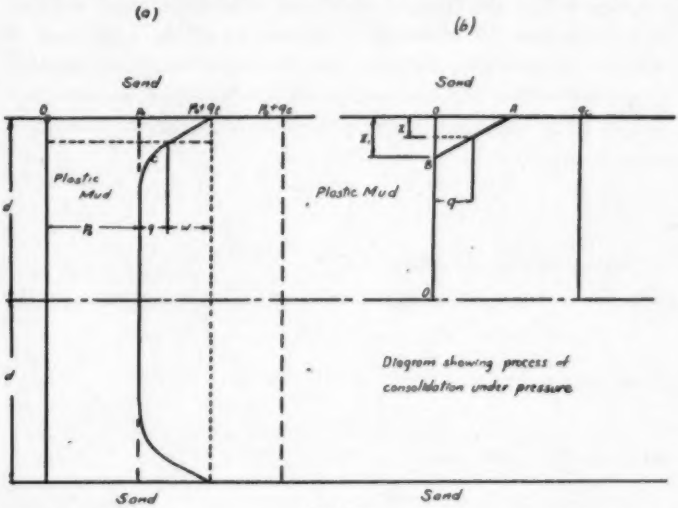


FIG. 3.

the hydrostatic pressure increases at an almost uniform rate until, at a certain depth z_1 beneath the surface, it becomes almost equal to q_t. From there on to the center its value remains practically constant and equal to q_t. As time goes on, both the values q_t and z_1 increase. The value of q_t can be expressed, according to Fig. 1*c*, by the formula

$$q_t = q_e \cdot \frac{t}{t_1} \cdot \qquad (1)$$

It remains to find out the law according to which z_1 increases.

For that purpose, we simplify the situation presented in Fig. 3*a* by replacing the curve C by a broken line ABC as shown in Fig. 3b.

As time goes on and consolidation proceeds, the point B, Fig. 3*b*, moves downward and the pressure q increases per unit of time by $\partial q/\partial t$.

According to our assumptions, an increase $\partial q/\partial t$ of the pressure q acting within the compressible layer causes the water content of the compressible material to decrease by $a \cdot \partial q/\partial t$ per unit of volume occupied by the solid substance. Hence the quantity Q of water which drains at a time t per unit of time and per unit of area of a horizontal plane at a depth z below the boundary plane is equal to

$$Q = a \int_z^{z_1} \frac{\partial q}{\partial t} \cdot dz. \tag{2}$$

According to Fig. 3*b* we have

$$q = q_t \frac{z_1 - z}{z_1} = q_t \left(1 - \frac{z}{z_1}\right).$$

Since both q_t and z_1 are variable, we obtain

$$\frac{\partial q}{\partial t} = \frac{\partial q_t}{\partial t}\left(1 - \frac{z}{z_1}\right) + q_t \frac{z}{z_1^2} \cdot \frac{dz_1}{dt}.$$

According to (1) we have

$$q_t = q_e \frac{t}{t_1}, \qquad \frac{q_t}{t} = \frac{q_e}{t_1},$$

hence

$$\frac{\partial q}{\partial t} = \frac{q_e}{t_1}\left(1 - \frac{z}{z_1}\right) + q_e \cdot \frac{t}{t_1} \cdot \frac{z}{z_1^2} \cdot \frac{dz_1}{dt}.$$

This introduced into (2) gives

$$Q = a \int_z^{z_1} \frac{q_e}{t_1}\left(1 - \frac{z}{z_1}\right) dz + a \int_z^{z_1} q_e \frac{t}{t_1} \frac{z}{z_1^2} \frac{dz_1}{dt} dz.$$

In connection with this operation, z is the only one variable,

while t and z_1 are supposed to remain constant. Hence

$$Q = \frac{aq_c}{t_1}\int_z^{z_1}\left(1 - \frac{z}{z_1}\right)dz + aq_c\frac{t}{t_1 z_1^2}\cdot\frac{dz_1}{dt}\int_z^{z_1} z\,dz$$

$$= \frac{a\cdot q_c}{t_1}\left[(z_1 - z) - \frac{1}{2z_1}(z_1^2 - z^2) + \frac{t}{2z_1^2}\cdot\frac{dz_1}{dt}(z_1^2 - z^2)\right] \qquad (3)$$

$$= \frac{aq_c}{t_1}\left[\frac{z_1}{2}\left(1 + \frac{t}{z_1}\frac{dz_1}{dt}\right) - z + \frac{z^2}{2}\left(\frac{1}{z_1} - \frac{t}{z_1^2}\cdot\frac{dz_1}{dt}\right)\right].$$

According to Darcy's law, the quantity of water which flows per unit of time through the unit of cross-section of a horizontal plane must be equal to the product of the coefficient of permeability k times the hydraulic gradient dw/dz. Hence

$$Q = k\cdot\frac{dw}{dz} \qquad \text{or} \qquad Qdz = kdw,$$

$$\int_0^{z_1} Qdz = k, \qquad \int_0^{z_1} dw = kq_t = kq_c\frac{t}{t_1}.$$

If we substitute for Q the term (3) we obtain

$$kq_c\frac{t}{t_1} = \frac{aq_c}{t_1}\left[\frac{z_1}{2}\left(1 + \frac{t}{z_1}\frac{dz_1}{dt}\right)\int_0^{z_1} dz - \int_0^{z_1} z\,dz + \frac{1}{2}\left(\frac{1}{z_1} - \frac{t}{z_1^2}\frac{dz_1}{dt}\right)\int_0^{z_1} z^2 dz\right]$$

$$= \frac{aq_c}{t_1}\left[\frac{z_1^2}{2}\left(1 + \frac{t}{z_1}\frac{dz_1}{dt}\right) - \frac{z_1^2}{2} + \frac{z_1^3}{6}\left(\frac{1}{z_1} - \frac{t}{z_1^2}\frac{dz_1}{dt}\right)\right]$$

$$= \frac{aq_c z_1}{6t_1}\left(z_1 + 2t\frac{dz_1}{dt}\right)$$

or

$$z_1\cdot\frac{dz_1}{dt} + \frac{z_1^2}{2t} = 3\frac{k}{a}.$$

The solution of this equation is

$$u = e^{-\int\frac{dt}{t}}\left(\int\frac{6k}{a}e^{\int\frac{dt}{t}}dt + C\right) = \frac{1}{t}\left(\frac{6k}{a}\int t\,dt + C\right)$$

$$= \frac{1}{t}\left[\frac{6kt^2}{2a} + C\right].$$

For $t = 0$ we have $z_1 = 0$, hence $C = 0$ and

$$z_1 = \sqrt{\frac{3kt}{a}}, \tag{4}$$

which means that the depth to which consolidation proceeds increases in direct proportion with the square root of time.

The settlement which occurred during the period 0 and t is obviously measured by the quantity of water which has left the layer during this period. The amount of water which flows through one boundary per unit of time is according to formula (2) equal to

$$Q_1 = a \int_0^{z_1} \frac{\partial q}{\partial t} dz = \frac{aq_e}{2t_1}\left(z_1 + t\frac{dz_1}{dt}\right).$$

According to formula (4) we have

$$z_1 = \sqrt{\frac{3k}{a}} \cdot \sqrt{t} \qquad \text{and} \qquad \frac{dz_1}{dt} = \sqrt{\frac{3k}{a}} \cdot \frac{1}{2\sqrt{t}}.$$

Hence

$$Q_1 = \frac{aq_e}{2t_1}\left(\sqrt{\frac{3kt}{a}} + \frac{1}{2}\sqrt{\frac{3kt}{a}}\right) = \frac{aq_e}{2t_1} \cdot \frac{3}{2}\sqrt{\frac{3kt}{a}} = \frac{3aq_e}{4t_1}\sqrt{\frac{3kt}{a}}.$$

The settlement s is equal to the thickness of the layer of water which has left the compressible stratum during the period 0 to t, through two boundaries. Hence

$$s = 2\int_0^t Q dt = \frac{3aq_e}{2t_1} \cdot \sqrt{\frac{3k}{a}} \int_0^t \sqrt{t}\, dt$$

$$= \frac{3aq_e}{2t_1} \cdot \sqrt{\frac{3k}{a}} \cdot \frac{2}{3}\frac{3}{t_2} = \frac{q_e}{t_1}\sqrt{3ka} \cdot t^{3/2}. \tag{5}$$

While consolidation proceeds, the point B, Fig. 3*b*, moves vertically downward towards the point 0. The time T at which point B arrives at 0 is determined by the relation (4);

$$z_1 = d = \frac{3kT}{a} \qquad \text{or} \qquad T = \frac{ad^2}{3k}. \tag{4a}$$

The further investigations depend on whether T is greater or smaller than the time t_1 at which the load acting on the surface of the ground becomes constant.

(1) *Case*: $T > t_1$.

As soon as the load becomes constant, further consolidation would proceed in the same manner as if the load had been constant since a certain time t_2. Hence we start with computing the time t at which full load should have been applied in order to produce at the time t_1 the settlement determined by formula 5, viz.,

$$s_1 = \frac{q_e}{t_1}\sqrt{3ka} \cdot t_1^{3/2} = \frac{q_e}{t_1}\sqrt{3kat_1}. \tag{5a}$$

In order to determine the relation between time and settlement on the assumption that full load had been applied at a time t_2, we start with equation (2):

$$Q = a\int_0^{z_1} \frac{dq}{dt} \cdot dz,$$

where in this time

$$q = q_e\left(1 - \frac{z}{z_1}\right), \qquad \frac{dq}{dt} - q_e \cdot \frac{z}{z_1^2} \cdot \frac{dz_1}{dt},$$

and

$$Q = a \cdot \frac{q_e}{z_1^2}\frac{dz_1}{dt}\int_0^{z_1} z dz = \frac{aq_e}{2} \cdot \frac{dz_1}{dt}.$$

It is easy to prove, that

$$z_1 = \sqrt{\frac{6k}{a}} \cdot \sqrt{t - t_2}, \tag{6}$$

and

$$\frac{dz_1}{dt} = \sqrt{\frac{6k}{a}} \cdot \frac{1}{2\sqrt{t - t_2}}.$$

Hence

$$Q = \frac{aq_e}{4} \cdot \sqrt{\frac{6k}{a}} \cdot \frac{1}{\sqrt{t - t_2}},$$

and

$$s = 2\int_{t_2}^{t} Q\,dt = \frac{aq_c}{2}\cdot\sqrt{\frac{6k}{a}}\cdot 2\sqrt{t - t_2} = q_c\sqrt{6ka(t - t_2)}. \quad (6b)$$

At $t = t_1$ the settlements should be equal to s_1. Hence

$$\frac{q_c}{t_1}\sqrt{3ka}\cdot t_1^{3/2} = q_c\sqrt{6ka(t_1 - t_2)},$$

or

$$t_2 = \frac{t_1}{2}. \quad (6c)$$

That means, *after the construction of the building is finished, the settlements would proceed according to the same law, as if the full load had been applied at once, at the time when construction was half finished.*

Further consolidation under constant load proceeds according to the formula

$$s = q_c\cdot\sqrt{6ka\left(t - \frac{t_1}{2}\right)}, \quad (6d)$$

which can be easily derived from equation (6), while the value z_1 increases according to the formula (6). If we introduce into this formula the values $z_1 = d$ and $t_2 = t_1/2$, we obtain for the time T at which $z_1 = d$ the expression

$$T = \frac{t_1}{2} + \frac{d^2a}{6k}.$$

From now on, z_1 remains constant and consolidation approximately proceeds as shown in Fig. 4*a*. In order to compute the further progress of consolidation we retain the symbols used in the preceding calculations. We further introduce

$$T' = T - \frac{t_1}{2} = \frac{d^2a}{6k},$$

and

$$t' = t - \frac{t_1}{2} \qquad dt' = dt.$$

From previous investigations [5] we know that for consolidation under constant load

$$q = q_e\left[1 - \frac{z}{d}\cdot e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)}\right].$$

According to formula (2), the quantity of water Q which flows per unit of time through the unit of area of one boundary is

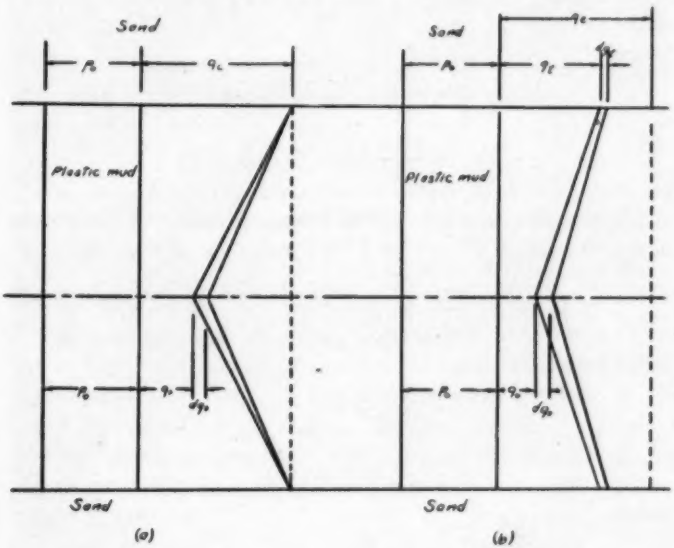


FIG. 4.

equal to

$$Q = a\int_0^d \frac{\partial q}{\partial t'}\cdot dz.$$

According to Fig. 4*a* we have

$$q = q_e - \frac{(q_e - q_0)z}{d},$$

and

$$\frac{\partial q}{\partial t} = \frac{z}{d} \cdot \frac{dq_0}{dt},$$

$$Q = \frac{a}{d}\frac{dq_0}{dt}\int_0^d z dz = \frac{a}{d}\frac{dq_0}{dt} \cdot \frac{d^2}{2} = \frac{ad}{2} \cdot \frac{dq_0}{dt}.$$

From formula (7) we obtain

$$\frac{dq_0}{dt} = +\frac{3kq_e}{ad^2} e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)},$$

hence

$$Q = \frac{3kq_e}{ad^2} \cdot \frac{ad}{2} \cdot e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)}$$

$$= \frac{3kq_e}{2d} \cdot e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)}.$$

The quantity of water which flowed through two boundaries during the period T' to $t' > T'$ is equal to

$$2\int_{T'}^{t'} Q dt = \frac{3kq_e}{d}\int_{T'}^{t'} e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)}, \qquad dt = \frac{3kq_e}{d} I.$$

Substitute in I:

$$-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right) = u,$$

$$-\frac{3k}{ad^2} dt = du,$$

Hence

$$I = -\frac{ad^2}{3k}\int e^u du = -\frac{ad^2}{3k} e^u$$

$$= -\left[\frac{ad^2}{3k} \cdot e^{\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)}\right]_{T'}^{t'},$$

and

$$2\int_{T'}^{t'} Q dt = -\frac{3kq_e}{d} \cdot \frac{ad^2}{3k}\left[e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)}\right]_{T'}^{t'}.$$

If we introduce into this formula our value

$$T' = \frac{d^2 a}{6k}$$

we obtain

$$s' = 2\int_{T'}^{t'} Q dt = adq_e \left[1 - e^{-\left(\frac{3k}{ad^2}t' - \frac{1}{2}\right)} \right]_{T'}^{t'}.$$

At the time T' the settlement was equal to $s_0 = q_e ad$. Hence, at the time $t' > T'$ the settlement will be equal to

$$s = q_e ad + s' = q_e ad \left\{ 2 - e^{-\left[\frac{3k}{ad^2}\left(t - \frac{t_1}{2}\right) - \frac{1}{2}\right]} \right\} \qquad (8)$$

For $t = \infty$ the second term in parenthesis becomes equal to zero, and the settlement becomes therefore equal to

$$s_e = 2q_e ad. \qquad (8a)$$

The settlement curve consists of three branches, extending over the time periods 0 to t_1, t_1 to T and T to ∞, determined by the equations (5), (6*d*) and (8)

(*2*) *Case*: $T < t_1$.

The preceding considerations were merely valid on the condition that z becomes equal to d after construction was finished, that means, after full dead and live load were applied. If this event occurs before construction was finished, that means for $T < t_1$, further consolidation will proceed according to Fig. 4*b*, but this time we have to deal with two variable pressures, viz., q_t and q_0. Hence

$$q = q_t - (q_t - q_0) \cdot \frac{z}{d}.$$

According to formula (1):

$$q_t = q_e \cdot \frac{t}{t_1}, \qquad \frac{dq_t}{dt} = \frac{q_e}{t_1} \qquad \text{and} \qquad \frac{\partial q}{\partial t} = \frac{dq_t}{dt} - \left(\frac{dq_t}{dt} - \frac{dq_0}{dt}\right)\frac{z}{d}.$$

Hence

$$\frac{\partial q}{\partial t} = \frac{q_e}{t_1} - \left(\frac{q_e}{t_1} - \frac{dq_0}{dt}\right)\frac{z}{d}.$$

According to formula (2) we have

$$Q = a\int_{d-z}^{d}\frac{\partial q}{\partial t}dz = a\left[\frac{q_e}{t_1}\int_{d-z}^{d}dz - \left(\frac{q_e}{t_1} - \frac{dq_e}{dt}\right)\frac{1}{d}\int_{d-z}^{d}zdz\right]$$
$$= a\left[\frac{q_e}{t_1}\cdot z - \left(\frac{q_e}{t_1} - \frac{dq_0}{dt}\right)\cdot\frac{1}{d}\cdot\frac{1}{2}(d^2 - d^2 + 2dz - z^2)\right]$$
$$= a\left[\frac{q_e}{t_1}z - \left(\frac{q_e}{t_1} - \frac{dq_0}{dt}\right)\cdot\frac{1}{2d}(2dz - z^2)\right]. \qquad (8a)$$

According to Darcy's law:

$$Qdz = kdw \text{ or } \int_0^d Qdz = k,\ \int_0^d dw = k(q_t - q_0) = k\left(q_e\frac{t}{t_1} - q_0\right),$$

$$\int_0^d Qdz = a\left[\frac{q_e}{t_1}\frac{d^2}{2} - \left(\frac{q_e}{t_1} - \frac{dq_0}{dt}\right)\cdot\frac{1}{2d}\left(\frac{2d^3}{2} - \frac{d^3}{3}\right)\right]$$
$$= a\left[\frac{q_e}{t_1}\cdot\frac{d^2}{2} - \left(\frac{q_e}{t_1} - \frac{dq_0}{dt}\right)\cdot\frac{d^2}{3}\right]$$
$$= a\left[\frac{q_e d^2}{6t_1} + \frac{d^2}{3}\cdot\frac{dq_0}{dt}\right].$$

Hence

$$k\left(q_e\frac{t}{t_1} - q_0\right) = \frac{ad^2}{3}\left(\frac{q_e}{2t_1} + \frac{dq_0}{dt}\right),$$

or

$$\frac{dq_0}{dt} - \frac{3kq_e}{ad^2t_1}\cdot t + \frac{3k}{ad^2}q_0 + \frac{q_e}{2t_1} = 0.$$

We substitute

$$\frac{3kq_e}{ad^2t_1} = A, \qquad \frac{3k}{ad^2} = B, \qquad \text{and} \qquad \frac{q_e}{2t_1} = D.$$

Hence

$$\frac{dq_0}{dt} - At + Bq_0 + D = 0.$$

Substitute

$$t = u + \frac{D}{A}, \qquad dt = du, \qquad u = t - \frac{D}{A},$$

$$\frac{dq_0}{du} - Au - D + Bq_0 + D = 0, \qquad \text{or} \qquad \frac{dq_0}{du} + Bq_0 = Au.$$

This is a linear equation whose integral is

$$q_0 = e^{-\frac{3k}{ad^2}\left(t - \frac{ad^2}{6k}\right)}\left[\frac{q_e a d^2}{kt_1}\left(\frac{kt}{ad^2} - \frac{1}{2}\right)e^{\frac{3k}{ad^2}\left(t - \frac{ad^2}{6k}\right)} + C\right],$$

or

$$q_0 = \frac{q_e a d^2}{kt_1}\left(\frac{kt}{ad^2} - \frac{1}{2}\right) + C \cdot e^{-\frac{3k}{ad^2}\left(t - \frac{ad^2}{6k}\right)}. \tag{9}$$

At a time T after construction was started, z_1 became equal to d and q_0 started with 0. From formula (4) we obtain

$$d = \sqrt{\frac{3kT}{a}}, \qquad \text{or} \qquad T = \frac{ad^2}{3k}. \tag{10}$$

If we introduce this value of T in (9) we obtain

$$0 = -\frac{q_e a d^2}{6kt_1} + C \cdot e^{-(1/2)}, \qquad \text{or} \qquad C = \frac{q_e a d^2}{6kt_1} \cdot e^{(1/2)}.$$

With this value of C, (9) passes into

$$\begin{aligned} q_0 &= \frac{q_e a d^2}{kt_1}\left(\frac{kt}{ad^2} - \frac{1}{2}\right) + \frac{q_e a d^2}{6kt_1} \cdot e^{-\left(\frac{3k}{ad^2}t - 1\right)} \\ &= \frac{q_e a d^2}{kt_1}\left[\frac{kt}{ad^2} - \frac{1}{2} + \frac{1}{6}e^{-\left(\frac{3k}{ad^2}t - 1\right)}\right] \\ &= \frac{q_e t}{t_1} - \frac{q_e a d^2}{2kt_1}\left[1 - \frac{1}{3}e^{-\left(\frac{3k}{ad^2}t - 1\right)}\right]. \end{aligned} \tag{10a}$$

From this term we obtain

$$\frac{dq_0}{dt} = \frac{q_e}{t_1} - \frac{q_e}{2t_1}e^{-\left(\frac{3k}{ad^2}t - 1\right)}.$$

The quantity of water Q which flows per unit of time through

the unit of area of one surface of the compressible layer can be obtained by introducing $z = d$ into (8a). Hence

$$Q = a\left[\frac{q_c d}{t_1} - \left(\frac{q_c}{t_1} - \frac{dq_c}{dt}\right)\frac{1}{2d}(2d^2 - d^2)\right]$$
$$= a\left(\frac{q_c d}{2t_1} + \frac{dq_0}{dt}\frac{d}{2}\right).$$

The value of dq_0/dt was obtained before. With this value we get

$$Q = a\left[\frac{q_c d}{2t_1} + \frac{q_c d}{2t_1} - \frac{q_c d}{4t_1} e^{-\left(\frac{3k}{ad^2}t - 1\right)}\right]$$
$$= \frac{aq_c d}{t_1}\left[1 - \frac{1}{4}e^{-\left(\frac{3k}{ad^2}t - 1\right)}\right].$$

The quantity of water which escapes during the period T to t through both top and bottom surface is

$$2\int_T^t Q dt = \frac{2aq_c d}{t_1}\left[t - T - \frac{1}{4}\int_T^t \underbrace{e^{-\left(\frac{3k}{ad^2}t - 1\right)}}_{T^1} dt\right].$$

To integrate T^1, substitute

$$-\left(\frac{3k}{ad^2}t - 1\right) = u, \qquad -\frac{3k}{ad^2}dt = du.$$

Hence

$$I = -\frac{ad^2}{3k}\int_T^t e^u du = -\frac{ad^2}{3k} \cdot e^{-\left(\frac{3k}{ad^2}t - 1\right)}\Big|_T^t.$$

From formula (4) we obtained

$$T = \frac{ad^2}{3k}.$$

For this reason

$$e^{-\left(\frac{3k}{ad^2}T - 1\right)} = 1$$

and

$$I = -\frac{ad^2}{3k}\left(e^{-\left(\frac{3kt}{ad^2} - 1\right)} - 1\right)$$

and

$$2\int_T^t Q\,dt = \frac{2aq_c d}{t_1}\left[t - T + \frac{1}{4}\frac{ad^2}{3k}\left(e^{-\left(\frac{3kt}{ad^2}-1\right)} - 1\right)\right]$$
$$= \frac{2aq_c d}{t_1}\left\{(t-T) - \frac{ad^2}{12k}\left[1 - e^{-\left(\frac{3kt}{ad^2}-1\right)}\right]\right\}.$$

Since the settlement is equal to the thickness of the layer of the water which escaped through the surface of the compressible layer, the last formula also represents the settlement which took place since the time T. Until the time T the settlement was equal to

$$s_1 = aq_c\frac{T}{t_1}d = q\cdot\frac{a^2d^3}{3t_1k}.$$

Therefore the settlement from $t = T$ to t (assuming $t_1 > t > T$) will be

$$s = q_c\frac{a^2d^3}{3t_1k} + q_c\frac{2ad}{t_1}\left\{t - T - \frac{ad^2}{12k}\left[1 - e^{-\left(\frac{3kt}{ad^2}-1\right)}\right]\right\}$$
$$= q\frac{a^2d^3}{6t_1k}\left[1 + \frac{12k}{ad^2}(t-T) + e^{-\left(\frac{3kt}{ad^2}-1\right)}\right]. \qquad (11)$$

At the time t_1 the settlement will be equal to

$$s_1 = q_c\frac{a^2d^3}{6t_1k}\left[1 + \frac{12k}{ad^2}(t_1 - T) + e^{-\left(\frac{3kt_1}{ad^2}-1\right)}\right]. \qquad (11a)$$

From now on, consolidation will proceed according to the laws which determine the process of consolidation under constant load. Since, in addition, the state $z_1 = d$ was already reached prior to t_1, the process will follow the same law from which formula (8) was obtained. In order to connect the process described by formula (11) with the process of consolidation under constant load, we have again to determine t_2 at which full load should be applied in order to produce at the time t_1 the settlement s_1, formula (11a). This condition requires according to formula (8)

$$s_1 = q_c ad\left\{2 - e^{-\left[\frac{3k}{ad^2}(t_1 - t_2) - \frac{1}{2}\right]}\right\}$$

from which we obtain

$$t_1 - t_2 = \frac{ad^2}{3k}\left[\frac{1}{2} - \log_e\left(2 - \frac{s_1}{adq_e}\right)\right]. \tag{12}$$

The remainder of the process of consolidation proceeds according to

$$s = q_e ad\left\{2 - e^{-\left[\frac{3k}{ad^2}(t - t_1) - \frac{1}{2}\right]}\right\}, \tag{13}$$

and for $t = \infty$ we obtain

$$s_e = 2q_e ad. \tag{14}$$

The settlement curve consists of three branches extending over the time periods 0 to T, T to t_1 and t_1 to ∞, determined by the equations (5), (11) and (13).

For reasons which are explained elsewhere,[5] the values k and d are so called "reduced values," which can be obtained from the true values (k_1 = true coefficient of permeability of the stratum, and $2d_1$ = true thickness of the stratum) by the following transformation:

$$k = \frac{k_1}{1 + n} \quad \text{and} \quad 2d = \frac{2d_1}{1 + n},$$

wherein n is the average voids ratio of the stratum (average between the voids ratio at the outset and at the end of the consolidation process).

The results of the preceding computations can be condensed into the following rules for determining the settlements during and after construction. Let

t_1	be the time when construction was completed,
n	the average voids ratio,
k_1	the true average coefficient of permeability,
$2d_1$	the true thickness of the compressible deposit,
$k = \frac{k_1}{1 + n}$	the reduced coefficient of permeability,

$2d = \dfrac{2d_1}{1+n}$ the reduced thickness of the deposit,

$T = \dfrac{ad^2}{3k}$ the time at which the point A, Fig. 1c, reaches the point 0, Fig. 1c,

q_e the vertical normal pressure which acts on the compressible layer after construction is finished,

s the compression of the layer after a time t,

s_e the compression of the layer after a time $t = \infty$, and

e the basis of the natural logarithms.

First case: $T > t_1$.

(a) $t = 0$ to $t = t_1$,

$$s = q_e \sqrt{\frac{3ka}{t_1}} \cdot t^{3/2}. \tag{5}$$

For

$$t = t_1, \qquad s = q_e \sqrt{3kat_1}.$$

(b) $t = t_1$ to $t = T$,

$$s = q_e \sqrt{6ka} \sqrt{t - \frac{t_1}{2}}. \tag{6d}$$

For

$$t = T, \qquad s = \sqrt{6ka} \sqrt{T - \frac{t_1}{2}}.$$

(c) $t = T$ to $t = \infty$,

$$s = q_e ad \left[2 - \frac{\sqrt{e}}{e^{\frac{3k}{ad^2}\left(t - \frac{t_1}{2}\right)}} \right]. \tag{8}$$

For

$$t = \infty, \qquad s_e = 2q_e ad.$$

Second Case: $T < t_1$.

(a) $t = 0$ to $t = T$,

$$s = q_e \cdot \frac{1}{t_1} \sqrt{3kat^{3/2}}. \tag{5}$$

For

$$t = T, \qquad s = q_c \frac{a^2 d^3}{3 t_1 k}.$$

(*b*) $t = T$ to $t = t_1$,

$$s = q_c \frac{a^2 d^3}{6 t_1 k}\left[1 + \frac{12k}{ad^2}(t - T) + \frac{e}{e^{\frac{3k}{ad^2}t}}\right]. \tag{11}$$

For

$$t = t_1, \qquad s_1 = q_c \frac{a^2 d^3}{6 t_1 k}\left[1 + \frac{12k}{ad^2}(t - T) + \frac{e}{e^{\frac{3k}{ad^2}t}}\right].$$

(*c*) $t = t_1$ to $t = \infty$,

$$s = q_c ad\left[2 - \frac{\sqrt{e}}{e^{\frac{3k}{ad^2}(t - t_2)}}\right], \tag{13}$$

wherein

$$t_2 = \frac{ad^2}{3k}\left[-\frac{1}{2} + \log_e\left(2 - \frac{s_1}{q_c ad}\right)\right] + t_1. \tag{12}$$

For

$$t = \infty, \qquad s_e = 2 q_c ad.$$

The following numerical computations should serve as an example. We assumed:

$t = 115$ feet $= 35$ m.,
$2d = 35$ feet $= 10.6$ m.,
$t_1 = 15$ months,
$q_c = 1$ kg. per cm.^{2}

We further assumed for the soil constants three different sets of values, assembled in table I. These sets were taken from actual test results. For the stiff material (*a*), the value of T was found to be greater than t_1 and for the two other ones smaller. Hence for the material (*a*) the computation of the settlements was performed by means of the formulae (5), (6d) and (8) (first case) and for the materials (*b*) and (*c*) by means

of the formulae (5), (11) and (13) (second case). The results of the computations are shown in Fig. 5. The abscissae represent the time in months and the ordinates the compression (settle-

TABLE I

Data for computing the process of consolidation for a compressible layer, 35 feet thick. Time $t_1 = 15$ months.

	(a) Fairly Stiff Clay	(b) Soft Clay	(c) Very Soft Clay
Original voids ratio n	1.83	2.02	2.40
Coeff. of compressibility a (g cm. system)	$0.137.10^{-3}$	$0.183.10^{-3}$	$0.292.10^{-3}$
True coefficient of permeability k_1 in cm. per min.	46.10^{-7}	114.10^{-7}	350.10^{-7}
Reduced coeff. of permeability k in cm. per min.	$16.7.10^{-7}$	$38.9.10^{-7}$	108.10^{-7}
Reduced coeff. of permeability in cm. per month	0.072	0.168	0.466
d in cm.	188	177	157
Ultimate compression $s_e = 2\ q_e ad$ in cm.	51.4	64.8	91.7
Ultimate compression s_e in feet	1.69	2.12	3.00
Time $T = ad^2/3k$ in months	22.5	11.4	5.2

ments) in centimeters. At the points A, B and C, which correspond to the time $t_1 = 15$ months, each curve shows a slight break. If full load had been applied at the time t_2, the settlements would have proceeded according to the dotted lines. Furthermore, if the compressible layer had been very permeable, such as to offer practically no resistance to the escape of the excess water, the settlements would have been as shown by the dash-dotted lines $OA'A''$, $OB'B''$ and $OC'C''$. The difference between the ordinates of these lines and the plain curves obviously represents the effect of the low permeability on the development of the settlements. The importance of this effect is greatest for the least permeable material (a) and smallest for the most permeable one (c). The Roman numbers I, II and III (Fig. 5) indicate those branches of the settlement curves, which correspond to the three equations which determine the shape of the curves (see page 20).

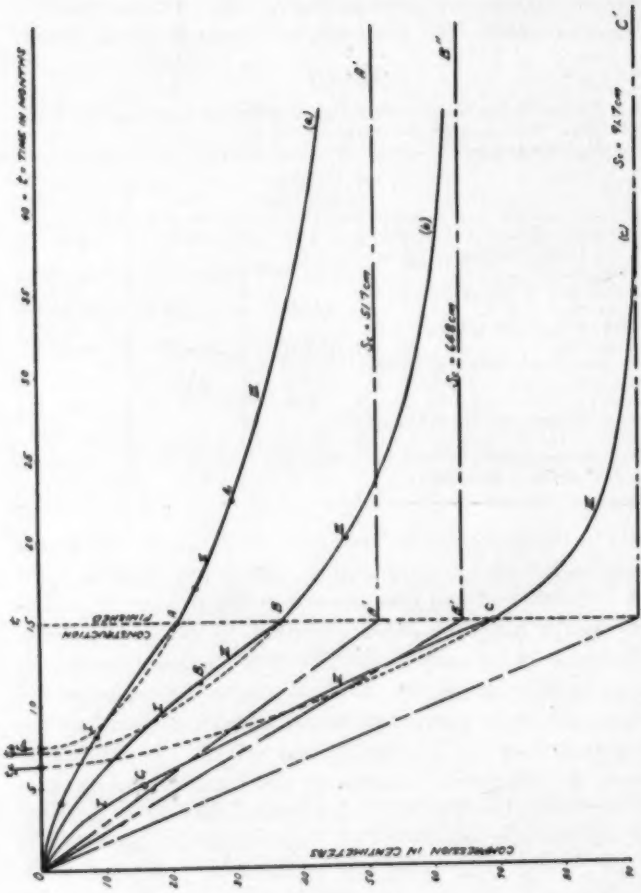

COMPRESSION IN CENTIMETERS
CONSTRUCTION FINISHED
(a)
(b)
(c)
$s_0 = 5.17$cm
$s_0 = 6.88$cm
$s_0 = 9.17$cm

FIG. 5.

In practice the time-settlement curves are almost never smooth as are the theoretical curves, but they have steps and other irregularities which do not reflect any change in the loading.

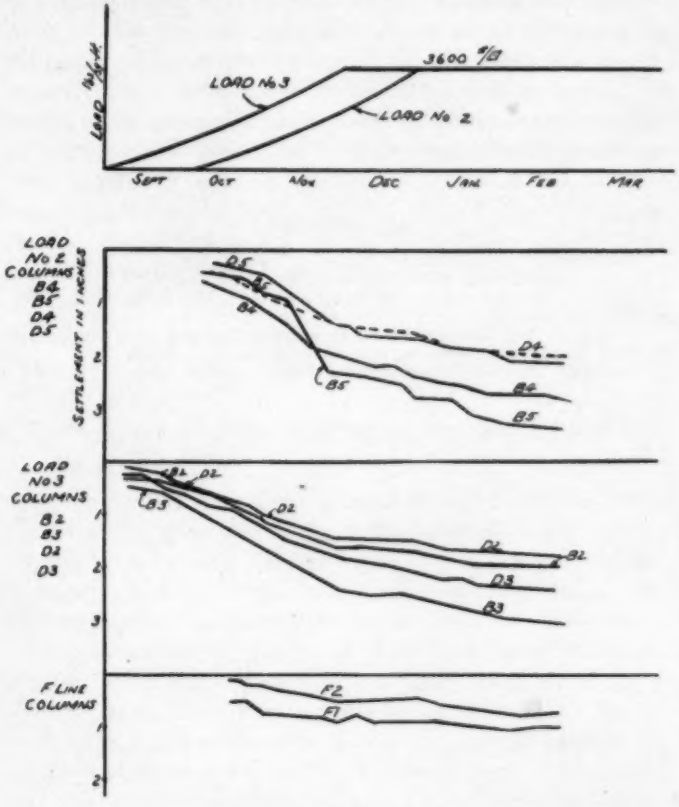


FIG. 6.

Thus, in the case Fig. 6 (settlement of large spread footings due to the gradual consolidation of a layer of soft clay), a very marked step of this kind extends over almost the whole month

of December. These steps seem to be connected with the stratification of the natural soil deposits. Theory is based among others on the assumption that the compressible layer is perfectly homogeneous. In non-homogeneous materials, part of the circulation of the water takes place through veins of more permeable material and the steps are suspected to be caused by the temporary obstruction of the openings or seams through which the excess water escapes. Yet, the general shape of the curves is similar to those in Fig. 5.

CAMBRIDGE, MASS.,
April 1929.

REFERENCES

[1] Boussinesq, "Application des Potentiels," Paris, 1885.
[2] Föppl, "Vorlesungen über Technische Mechanik."
[3] Kögler, "Druckverteilung im Baugrund. II." Die Bautechnik, 1928, pp. 205–209, 229–232.
[4] Terzaghi, "Soil Classification for Foundation Purposes." Proceedings of the First International Soil Congress, Washington, D. C., 1928.
[5] Terzaghi, "Erdbaumechanik," Wien, 1925.

TESTS ON HIGH-PRESSURE PIPE BENDS

By William Hovgaard

A. Introduction

This paper contains a report and analysis of a series of tests on high-pressure pipe bends, supplementing tests previously made on pipes designed for moderate steam pressures. It is the fourth paper by the author on this subject published in the Journal of Mathematics and Physics of the Massachusetts Institute of Technology, being preceded by the following:

First paper: "The Elastic Deformation of Pipe Bends," Vol. VI, No. 2, Nov. 1926.

Second paper: "Deformation of Plane Pipes," Vol. VII, No. 3, Oct. 1928.

Third paper: "Further Research on Pipe Bends," Vol. VII, No. 4, Dec. 1928.

In those papers were given formulas for calculating the deflection of a pipe bend subject to certain external forces, or, conversely, for calculating the forces created in a pipe bend subject to heat expansion in a pipe line of which the bend forms part. The question of the stresses arising in pipe bends under these conditions and hence the strength of the bends under the action of given forces and couples was also fairly cleared up as far as pipes designed for moderate pressures are concerned. In the third paper it was proposed to use the longitudinal stress in the middle surface as the criterion of strength, and it was found that so long as that stress did not exceed about 20,000 lbs./sq. in. in pipes made of soft steel, lap-welded or seamless, there would be no appreciable permanent set in the bend as a whole, although local overstrains might exist over limited areas of the surface of the pipe wall in the neighborhood of the most strained section.

It was not considered safe, however, to rely on this strength criterion for pipes of relatively greater wall thickness until it had been verified experimentally. The tests described in the present paper were made chiefly for the purpose of clearing up this point.

In previous tests the longitudinal strains were obtained by means of Berry strain gages, which were applied by hand successively to holes drilled two inches apart longitudinally at equal intervals around the circumference of the most strained section. It was found that the strains so obtained were not always accurate, in some cases even erratic, probably on account of the difficulty of placing the gages exactly in the same position in the holes each time. In order to obviate this difficulty nine Huggenberger tensometers* were acquired and were used in the tests of two of the bends. These instruments seemed to give more consistent results and, being fixed in their positions during the test, did not require resetting for each reading. At any time the strain could be easily ascertained and if a local overstrain or flow of the material took place, it often became visible, after the load had become stationary, by a continued movement of the pointer of the respective instrument. This movement continued until an adjustment between the plastic and elastic strains had taken place whereby equilibrium between the external bending moment and the stress couple was again established.

In order to enable the reader to view the tests here recorded in the right perspective, a list is given in Table 1 of all the tests made on this subject at the Massachusetts Institute of Technology, together with some of the principal characteristics of the respective pipes. The symbols used in this Table and in the following have the same meaning as in previous papers, but for convenience they will be here defined again.

P = Force acting at each end of the pipe bend

M = Bending moment

* Hereafter often referred to as the "tensometers" while the Berry strain gages are referred to as the "Berry gages."

R = Radius of curvature of axis of pipe

$t = 2\ h$ = Wall thickness

r = Radius of section of middle surface of pipe wall, assuming this section to be a circle. When the section deviates from the circular form, r is the radius of curvature.

TABLE 1

No.	Origin of Pipe Bend	Nominal Diam. Inches	When tested	$\frac{2r}{t}$	$\frac{R}{2r}$	K	λ	Shape
1	Boston Navy Yard	4½	1926	39.2	3.0	5.24	.304	
2	Walworth Mfg. Co.	6	1926	24.2	4.7	2.07	.785	
3	Boston Navy Yard	9½	1928	36.5	4.6	3.19	.509	
4	Boston Navy Yard	14	1928	35.6	5.9	2.44	.662	
5	Walworth Mfg. Co.	6	1928	24.2	4.7	2.07	.785	
6	National Tube Co.	8	1929	15.2	4.4	1.53	1.159	
7	National Tube Co.	10	1929	19.5	3.2	1.61	1.074	
8	National Tube Co.	12	1929	27.8	4.8	2.32	.696	
9	Walworth, Stone & Webster	6	1929	18.0	4.9	1.59	1.090	

The following two non-dimensional coefficients are used:

$$K = \frac{48h^2R^2 + 10r^4}{48h^2R^2 + r^4}, \tag{1}$$

$$\lambda = \frac{2hR}{r^2}, \tag{2}$$

$$K = \frac{12\lambda^2 + 10}{12\lambda^2 + 1}, \tag{1'}$$

$$\lambda^2 = \frac{10 - K}{12(K - 1)}. \tag{2'}$$

Bends Nos. 2 and 5 were actually identical, but are given separately because test No. 5 was carried out after the pipe had been slightly overstrained in test No. 2, so that it might be regarded virtually as a new pipe with somewhat different characteristics.

Bends Nos. 1, 3 and 4 were of the thin-walled type, the thickness being not much more than one-fortieth of the diameter of the middle surface. Bends Nos. 2, 5 and 8 were of moderate wall thickness, suitable for medium steam pressures, but bends Nos. 6, 7 and 9 were made from high-pressure pipes, having thicknesses of more than one-twentieth of the diameter.

The radius of curvature R to which the pipes were bent at the top of the curve was less than five times the diameter, except in case of the 14-in. pipe, No. 4, where this ratio was about six.

Where convenient we shall refer to the pipe bends and to the tests by their number as given in Table 1.

B. Description of the Pipe Bends Recently Tested

The bends to be described here are numbered 6, 7, 8 and 9 in Table 1.

The pipes Nos. 6, 7 and 8, which were of respectively 8 in., 10 in., and 12 in. nominal diameter, were presented to the Institute in the straight condition by the National Tube Company of Pittsburgh for the purpose of experimentation. All three pipes were lap-welded steel tubes. They were bent hot and provided with flanges in the Boston Navy Yard. The data for tensile tests and chemical analysis furnished by the manufacturers are giyen in Table 2, which shows that the material was a soft steel of good quality.

The modulus of elasticity for the three pipes was taken $E = 30 \times 10^6$ lbs./sq. in.

Figs. 1, 2 and 3 give the general form and dimensions of these bends.

The pipe bend No. 9 of 6 inch nominal diameter was presented to the Institute by the Stone & Webster Engineering

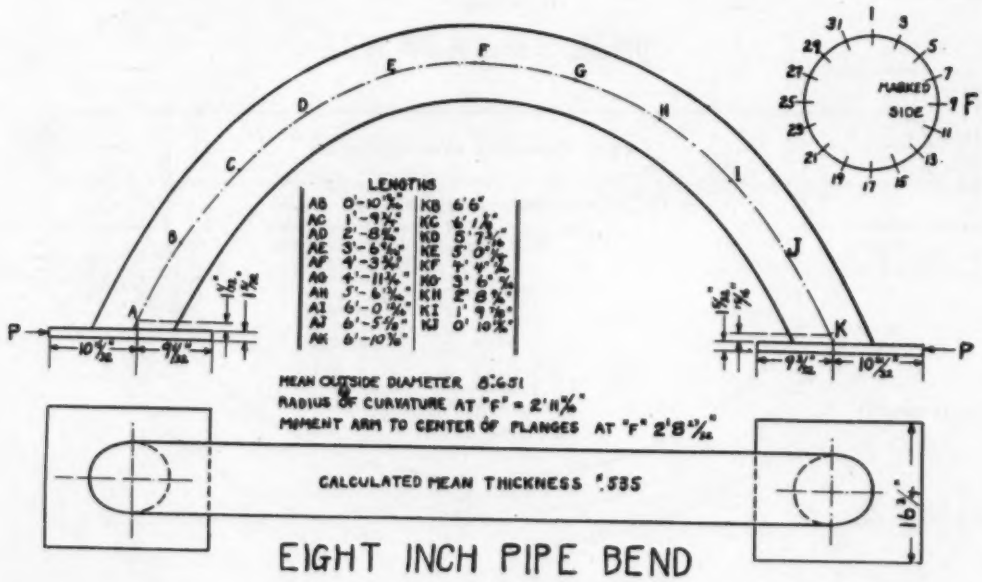
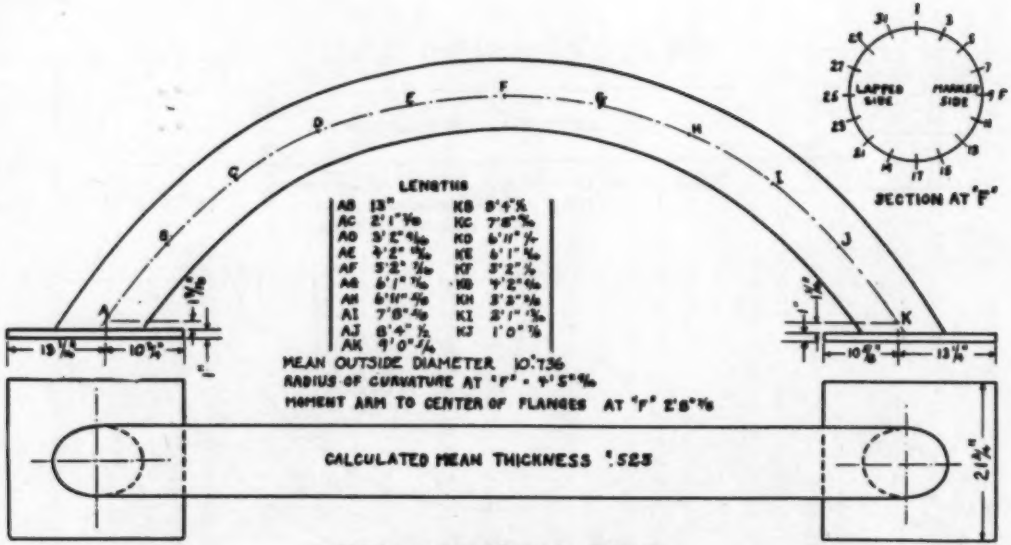
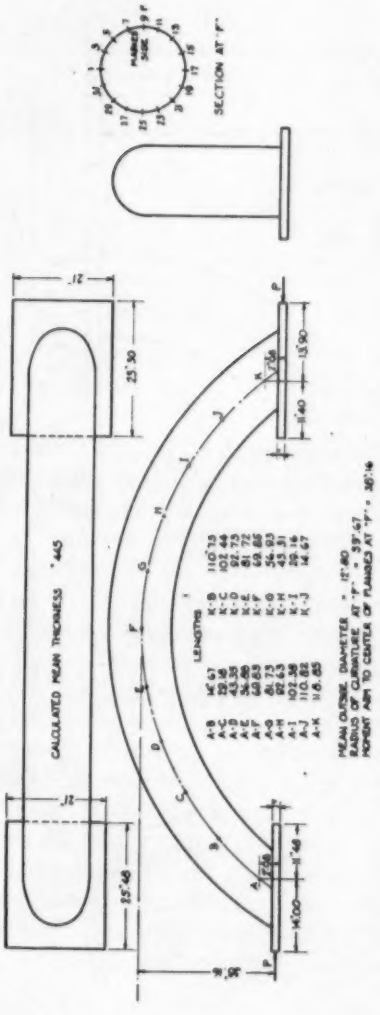

LENGTHS
AB 0'-10 7/16"
AC 1'-9 1/16"
AD 2'-8 5/16"
AE 3'-6 9/16"
AF 4'-3 3/16"
AG 4'-11 1/4"
AH 5'-6 1/16"
AI 6'-0 7/16"
AJ 6'-5 1/16"
AK 6'-10 7/16"
KB 6'6"
KC 6'1 1/16"
KD 5'7 3/16"
KE 5'0 7/16"
KF 4'4 9/16"
KG 3'6 5/16"
KH 2'8 1/16"
KI 1'9 7/16"
KJ 0'10 7/16"
MARKED SIDE
MEAN OUTSIDE DIAMETER 8.651
RADIUS OF CURVATURE AT "F" = 2'11 7/16"
MOMENT ARM TO CENTER OF FLANGES AT "F" 2'8 1/16"
CALCULATED MEAN THICKNESS .535
16 1/4"
EIGHT INCH PIPE BEND

FIG. 1.



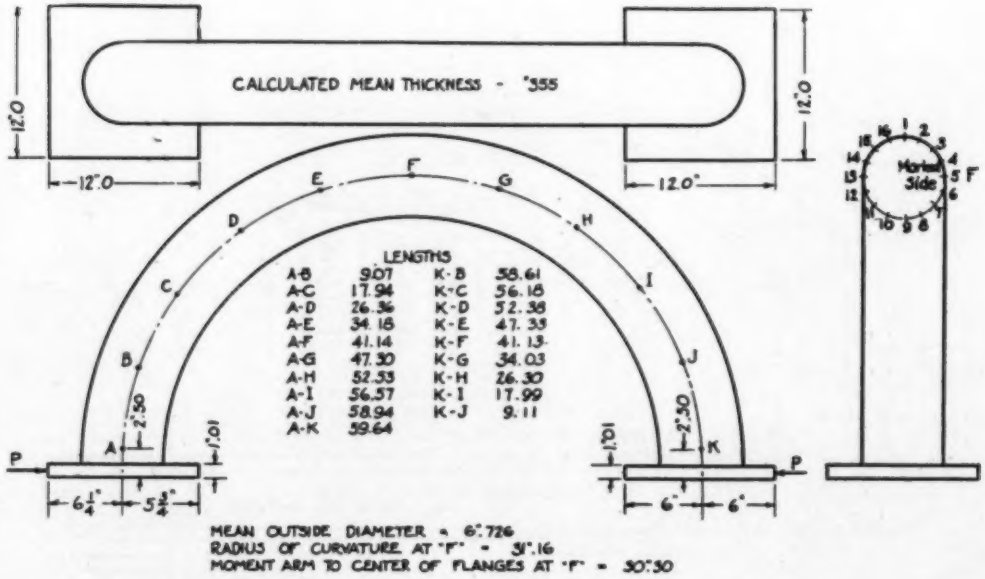
TEN INCH PIPE BEND

FIG. 2.


SECTION AT 'F'
CALCULATED MEAN THICKNESS '445
LENGTHS '
MEAN OUTSIDE DIAMETER = 12'40
RADIUS OF CURVATURE AT 'F' = 39'67
RADIUS FROM TO CENTRE OF FLANGES AT 'F' = 38'14

TWELVE INCH PIPE BEND

FIG. 3.



SIX INCH PIPE BEND

FIG. 4.

TABLE 2

	Diameter of Pipe	8 in.	10 in.	12 in.
Tensile Properties	Yield point lbs./sq. in.	27,130	29,830	29,730
	Tensile strength lbs./sq. in.	47,000	54,230	49,340
	Elongation in 8 inches, per cent	36.0	35.0	35.0
Chemical Analysis	Sulphur, per cent	.035	.047	.043
	Phosphorus, per cent	.023	.032	.013
	Manganese, per cent	.43	.61	.45
	Carbon, per cent	.10	.11	.10

Corporation and the Walworth Manufacturing Company. The latter fabricated the bend from seamless steel tubing manufactured by the Pittsburgh Steel Products Company. The bend was delivered with two square flange plates, machined at the outer edges, ready to be set up in the testing machine. The characteristics of the material were determined at the Institute by compressing a straight 30 inch length of the same pipe from which the bend was made. After the ends had been machined square, the pipe was inserted in a testing machine and subjected to simple compression. It was found that the elastic limit was 42,700 lbs./sq. in. and the modulus of elasticity, $E = 29{,}150{,}000$ lbs./sq. in., but judging from the behavior of the bend under the tests, it appears that the yield point was much lower in the bent pipe, probably due to the heat treatment to which it was subjected in the process of bending.

Fig. 4 gives the general form and dimensions of this bend. It will be noticed that while pipes Nos. 6, 7 and 8 were bent approximately to a quadrant, pipe No. 9 was bent to a complete semi-circle.

The method of setting up the bends in the testing machine and the use of knife edges at the edge of the flanges in order to

avoid terminal couples was the same as in previous tests and is explained in the first and third papers. See in particular Fig. 3 of the last named paper.

C. Description of the Tests

Measurements with strain meters and micrometer calipers were taken around the section *F* at the top of the bend as in previous tests.

In pipes Nos. 6, 7 and 8, section *F* was divided into 32 equal parts as indicated in Figs. 1, 2 and 3, and at each division two holes were drilled, spaced 2 inches apart in a longitudinal line, the holes being just deep enough to accommodate the conical points of the Berry strain gages. At every alternate division the surface was polished for caliper measurements.

While the strain readings on pipes Nos. 6 and 7 were obtained entirely by means of Berry gages, those on No. 8 and No. 9 were obtained partly by the Huggenberger instruments.

The 8-in. Pipe Bend.—An initial load of 1500 lbs. was applied in order to eliminate the effects of small irregular adjustments at the face plates and the knife edges, and the readings obtained at this load were used as a basis of reference for all subsequent readings, so that the strains corresponded not to the actual or gross load on the bend, but to the net load or difference between the actual and the initial load. Ultimately the stresses obtained on this basis were augmented proportionately so as to correspond to the gross load.

After a complete set of readings had been taken in the initial condition, the load was increased to 12,000 lbs. and readings were taken again. According to a preliminary calculation the stresses should still be within the elastic limit.

The load was now increased by increments of 1000 lbs. Soon a slight cracking of scale began to be heard and increased gradually, but the rate of approach of the ends of the pipe to each other, that is, the decrements in the distance *AK* in Fig. 1, remained proportional to the increments in the load.

At a load of 24,000 lbs. the cracking of scale became more pronounced and a few well-defined strain lines appeared at the top of the bend, extending at first from section *F* and half way to *G*. Transversely these lines extended across the upper part of the pipe between points 6 and 30. After the test the lines extended through *F* in both directions beyond *D* and *H* and appeared also on the lower side of the pipe. See Figs. 23 and 24.

The rate of shortening of *AK*—the distance between the end marks—remained proportional to the increase in load up to 23,000 lbs., but when the load was raised to 24,000 lbs. the decrement increased. It was clear that flow of the material had occurred locally and that failure was imminent. After a complete set of readings had been taken, the load was reduced to 1500 lbs. and it was then found that there was a permanent deflection of $\frac{5}{16}$ inch.

The load was again raised to 24,000 lbs., which the pipe seemed capable of holding indefinitely, but a further increment of 1000 lbs. showed clearly that the limit of proportionality had been passed, and when it was attempted to increase the load further, the beam of the testing machine dropped at 25,700 lbs. and the load began to fall off while the deflection increased steadily. At this point, which is referred to as that of "complete breakdown," the elastic strength of the pipe seemed to be completely exhausted at the most strained section and the load was taken off. During the period of breakdown cracking of scale proceeded and the strain lines became more pronounced and extended.

It is of interest to note the suddenness with which breakdown took place. Up to about 23,000 lbs. the pipe as a whole seemed to preserve its elasticity, but within a range of less than 3000 lbs. complete failure occurred. Actually the breakdown took place gradually by a local flow of the metal in certain regions of the most strained section, and the plastic areas so formed were progressively extended, augmenting the strains in the surrounding elastic material.

The 10-in. Pipe Bend.—This bend was tested in the same

manner as the 8-in. bend. Initial readings were taken at a load of 2000 lbs. and the next complete set of readings at 18,000 lbs. After that the load was increased stepwise to 39,000 lbs., when again a complete set of readings was taken.

Up to a load of 36,000 lbs. there seemed to be perfect proportionality between the decrements of AK and the increments in the load, but from 36,000 lbs. to 39,000 lbs. a progressive increase in the deflection took place.

Next the load was increased stepwise to 41,000 lbs. but at this load a creeping in the deflection took place, attaining a rate of about 1 inch per minute. About ten minutes after the load had reached 41,000 lbs. the load began to fall off and in two minutes dropped to 36,000 lbs., showing that complete breakdown had occurred and the load was taken off. The final deflection, that is, the shortening of AK, was then found to be $5^{5}/_{32}$ inches.

Strain lines were first observed at 39,000 lbs., and extended finally across the top of the pipe from sections C–D to H–I from points 5–6 to 29, all inclined at 45 degrees to the axis. They were also found on the lower side as in the 8-in. pipe.

The 12-in. Pipe Bend.—In this test Huggenberger tensometers were used, being applied at the nine odd numbered points 1 to 17 on the marked side of the pipe. At all other points Berry gages were used. The method of attaching the tensometers did not prove quite satisfactory and a better method was devised for the 6-in. pipe as described below.

The initial or reference load was 5000 lbs. and the next load at which a complete set of measurements was taken was 18,000 lbs. where the stresses were still well within the elastic limit. After that the load was increased stepwise to 30,000 lbs., at which point scale began to fly off and readings were taken. Proportionality of the deflections existed up to this point and the pipe was able to hold the load, which was probably the highest safe load. As the load was increased beyond 30,000 lbs., proportionality of the deflections ceased and complete break-

down took place at 34,500 lbs. Diagonal strain lines were found on the upper part of the bend on both sides of the longitudinal middle line, beginning at about points 2 and 31. In the region between the diagonal lines and extending from F towards the ends beyond E and G, were formed a bunch of longitudinal lines, as seen on the picture, Fig. 25. On the lower side of the pipe the lines were less pronounced and regular than on the upper side.

Again, as in the two previous tests, the pipe as a whole passed from a condition of apparent elasticity to complete breakdown within a range of loading of a few thousand pounds. After the load had been taken off, the total shortening in AK was found to be 2.4 inches.

The 6-in. Pipe Bend.—In this case the section at F, as seen in Fig. 4, was divided into only 16 intervals on account of the small diameter of the pipe. Huggenberger tensometers were applied at the eight odd-numbered points, while Berry gages were used at the intermediate even-numbered points. The fixing devices and the general set-up of the tensometers can be seen in Fig. 5. Each instrument was clamped down by an elastic wire bridge, which was pressed in towards the pipe by a screw entering a tapped hole in the pipe wall. The holes were of about one-eighth of an inch diameter and were at a distance of about five inches from the section F where the strains were measured. This method of holding the gages proved satisfactory and did not interfere with the use of Berry gages or calipers at intermediate points.

Before the tensometers were attached, a complete set of longitudinal strain measurements were made with Berry gages at all sixteen points, first at zero load and then at the initial load which in this case was 500 lbs. Then the tensometers were fitted in place and readings on them were taken.

Henceforth all complete sets of readings were obtained on the tensometers at odd-numbered points and on the Berry gages at even-numbered points. Such readings were taken at loads of 2500, 5000, 6500, 6700, 8500, 10,500 and 16,500 lbs. Moreover,

after each of the loads 2500, 5000, 6500 and 6700 lbs. had been applied, the load was reduced to 500 lbs. and readings on the tensometers were taken.

FIG. 5. Six-Inch Pipe Bend. Attachment of Huggenberger Tensometers.

The shortening of AK was proportional to the increase in load up to 8500 lbs., the decrement being from .030 to .035 of an

inch for 500 lbs. increase in load. At 6500 lbs. it was observed that some of the tensometers continued to move for a short time immediately after the load had become stationary, yet, while this load was kept on for 30 minutes, no creeping in the length AK was noticeable.

From 8500 to 12,000 lbs. load, the decrement in AK was .040 inch for each 500 lbs., but after that it increased more rapidly. When 16,500 lbs. was held, a slight creeping reduction in AK was observed, amounting to .04 inches in four minutes, but then the creeping ceased and equilibrium was established.

From 16,500 lbs. the load was again augmented in steps of 500 lbs., but when 19,200 lbs. was reached, the load began to fall off under steadily increasing deflection of the pipe and the load was then taken off. At zero load and after the strain meters had been removed a complete set of readings were again taken with Berry gages.

No strain lines were visible on the 6-in. pipe, being probably concealed by the protective paint with which it was coated.

The breakdown did not actually take place at the F section but more nearly at E where the pipe showed a pronounced flattening after the test.

D. Analysis of the Tests

Calculation and Plotting of the Stresses and Strains.—The measurements obtained from the Berry and Huggenberger instruments gave directly the longitudinal strains, ϵ_{01}, while the caliper measurements around the transverse section furnished the means of computing the transverse strains, ϵ_{02}. As explained in the third paper the radius of curvature at any point of the transverse section was calculated from the formula:

$$r = a^2b^2\rho^3\left[\frac{\cos^2\theta}{a^4} + \frac{\sin^2\theta}{b^4}\right]^{3/2}, \qquad (3)$$

where a and b are the semi-axes of the transverse section, which is assumed to be of elliptic form; ρ is the radius vector at any

angle θ referred to the radius normal to the plane of the pipe bend as initial line, OA in Fig. 6.

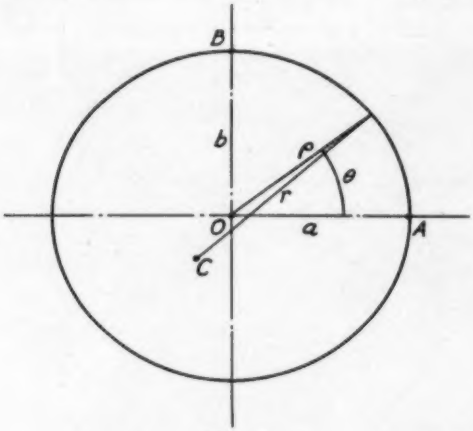


FIG. 6.

Let r and r' be the radii of curvature before and after deformation, then the transverse strain at any point in the pipe wall, distant z from the middle surface is approximately:

$$\epsilon_{02} = z\left(\frac{1}{r'} - \frac{1}{r}\right). \tag{4}$$

The longitudinal and transverse "observed" stresses, p_{01} and p_{02} respectively, were computed from the formulas:

$$p_{01} = \frac{m^2 E}{m^2 - 1}\left[\epsilon_{01} + \frac{\epsilon_{02}}{m}\right], \tag{5}$$

$$p_{02} = \frac{m^2 E}{m^2 - 1}\left[\epsilon_{02} + \frac{\epsilon_{01}}{m}\right], \tag{6}$$

where m is Poisson's number, here taken equal to 10/3.

These stresses were compared with the "calculated" stresses p_1 and p_2 obtained by a purely theoretical calculation from the

following formulas, the derivation of which is explained in the third paper:

$$p_1 = \frac{KMr}{I}\left[\sin\theta - \frac{6r^4\sin^3\theta}{24h^2R^2 + 5r^4} + \frac{18hr^2R\cos 2\theta}{m(24h^2R^2 + 5r^4)}\right] + p_c, \quad (7)$$

$$p_2 = \frac{KMr}{I}\left[\frac{\sin\theta}{m} - \frac{6r^4\sin^3\theta}{m(24h^2R^2 + 5r^4)} + \frac{18hr^2R\cos 2\theta}{24h^2R^2 + 5r^4}\right] \quad (8)$$

where r is now the radius of the section of the middle surface, assumed to be a circle. The value of r is obtained as one-half the difference between the average outside diameter at F minus the wall thickness. I is the moment of inertia about the neutral axis of the area of the transverse section of the pipe wall. The additional term, p_c, in (7)is the longitudinal compressive stress due to the end load, which at the top section is found simply from the formula:

$$p_c = \frac{P}{A}, \quad (9)$$

where A is the sectional area of the pipe wall.

At the middle surface the transverse stress is assumed to be zero and the longitudinal stress is found from:

$$p_{m1} = \frac{KMr}{I}\left[\sin\theta - \frac{6r^4\sin^3\theta}{24h^2R^2 + 5r^4}\right] + p_c, \quad (10)$$

the maximum value of which, for ordinary values of K, is

$$p_{m1,\,max.} = \pm\frac{2}{3}\frac{KM}{rI}\sqrt{\frac{24h^2R^2 + 5r^4}{18}} + p_c, \quad (11)$$

which may also be written

$$p_{m1,\,max.} = \pm\frac{2}{3}\frac{KMr}{I}\sqrt{\frac{6\lambda^2 + 5}{18}} + p_c = \frac{KMr}{3I}\sqrt{\frac{K}{K-1}} + p_c \quad (11')$$

If $K \lessapprox 1.333$ or $\lambda \gtrsim 1.472$, the maximum stress is found by making $\sin\theta = 1$ in (10) as explained below.

Often in practice p_c is so small that it can be neglected, in

which case the stress at the middle surface is of the same magnitude in tension as in compression.

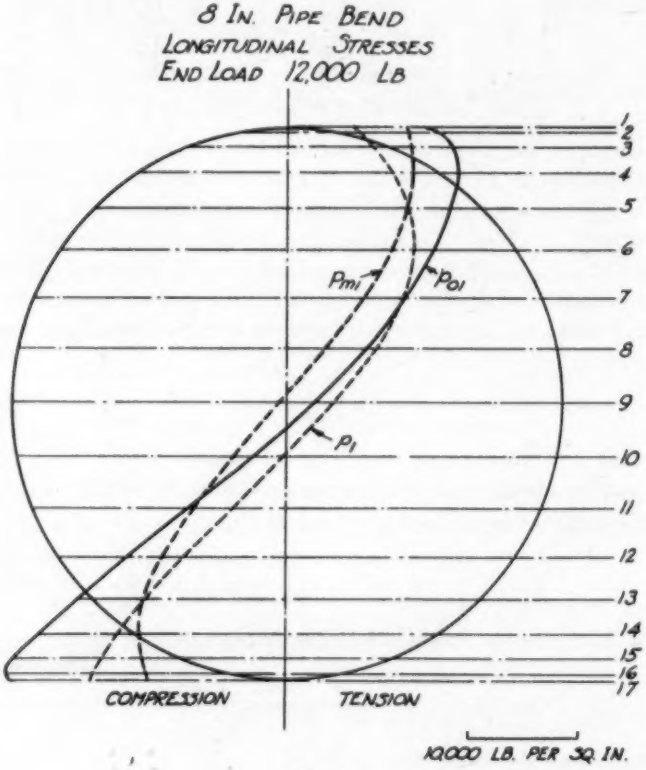


FIG. 7.

By means of the above formulas the stresses were calculated for each of the four pipes for various loads and plotted in diagrams in the same manner as in the first and third paper.

The diagram for the 8-in. pipe, Fig. 7, gives curves for the calculated longitudinal stresses for the outer and the middle

surface, p_1 and p_{m1}, and for the observed longitudinal stresses for the outer surface, p_{01}, all for an actual or total load, $P = 12{,}000$ lbs. Fig. 8 shows the calculated and observed transverse stresses for the 8-in. pipe at the same load.

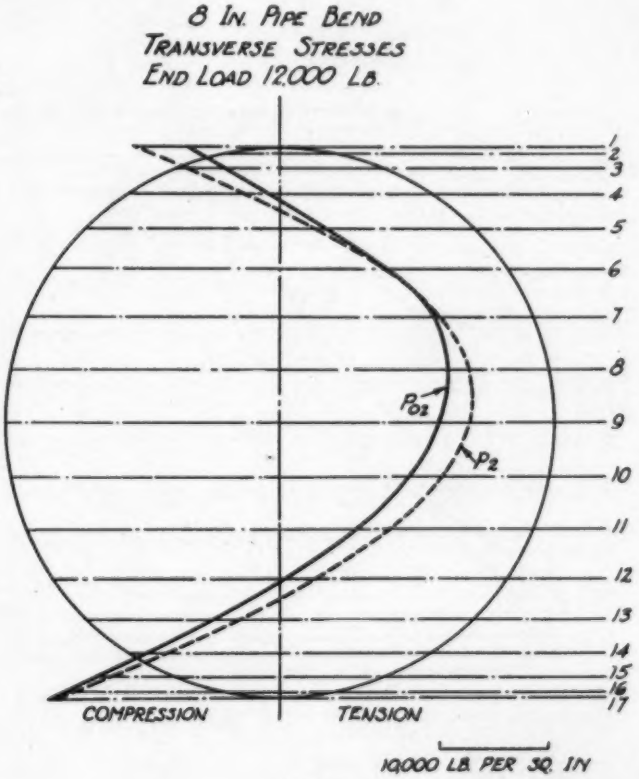


FIG. 8.

So also Figs. 9 and 10 give similar curves of stress for the 10-in. pipe at a load of 18,000 lbs., and Figs. 11 and 12 for the 12-in. pipe at a load of 18,000 lbs.

The curves give the stresses only for one side of the pipe section. The calculated stresses are of course the same on both sides, but in case of the "observed" stresses, which are usually

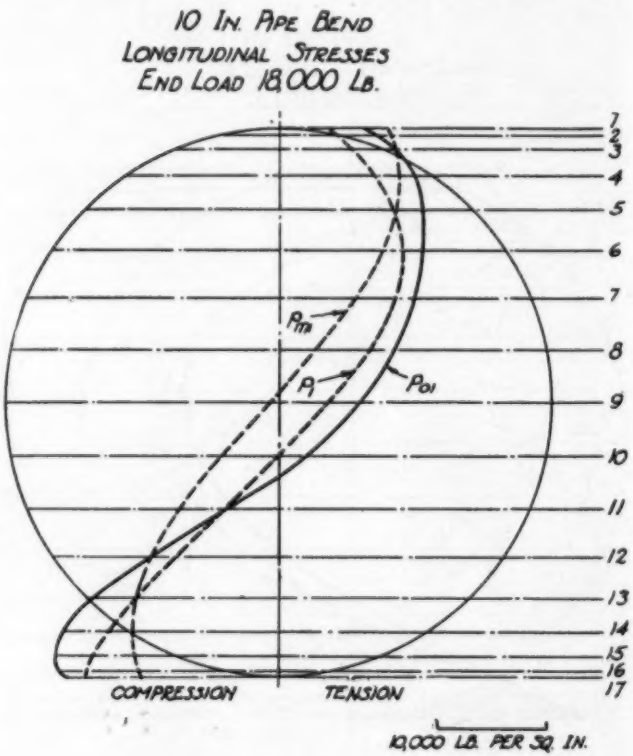


FIG. 9.

different on the two sides, the mean values were plotted. The observed stresses, being first obtained from strains corresponding to the "net" load, were augmented before plotting so as to correspond to the actual total or gross load. Thus, for instance, the

values of p_{01} and p_{02} for the 8-in. pipe for a gross load of 12,000 lbs. were calculated first on the basis of the strains ϵ_{01} and ϵ_{02} corresponding to a load of 12,000 − 1500 = 10,500 lbs., and then augmented in the ratio 12,000/10,500.

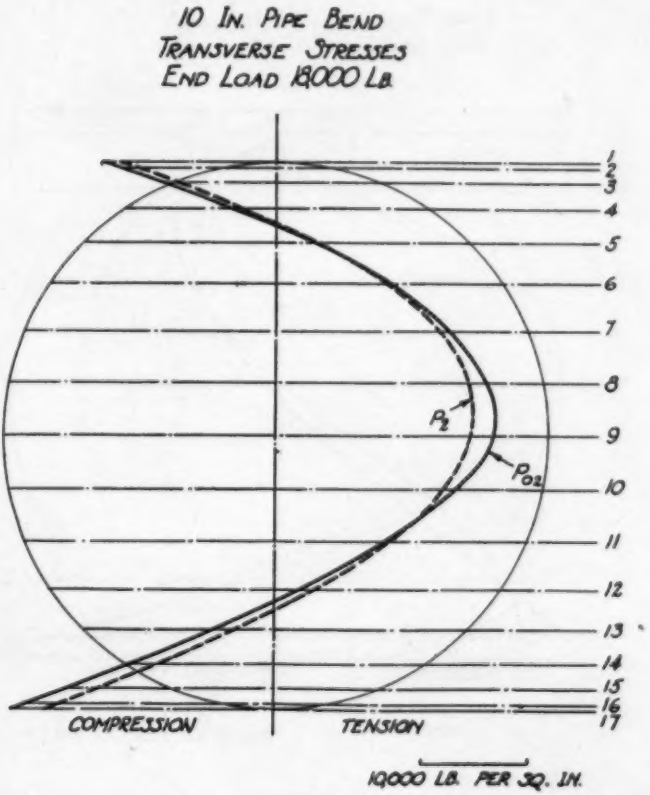


FIG. 10.

Figs. 13 to 18 give the stress diagrams for the 6-in. pipe for loads of 2500, 5000 and 6500 pounds. The observed strains

were here obtained as the difference between the readings at those loads and the readings at a load of 500 lbs. applied immediately after, while in previous tests the readings at the original

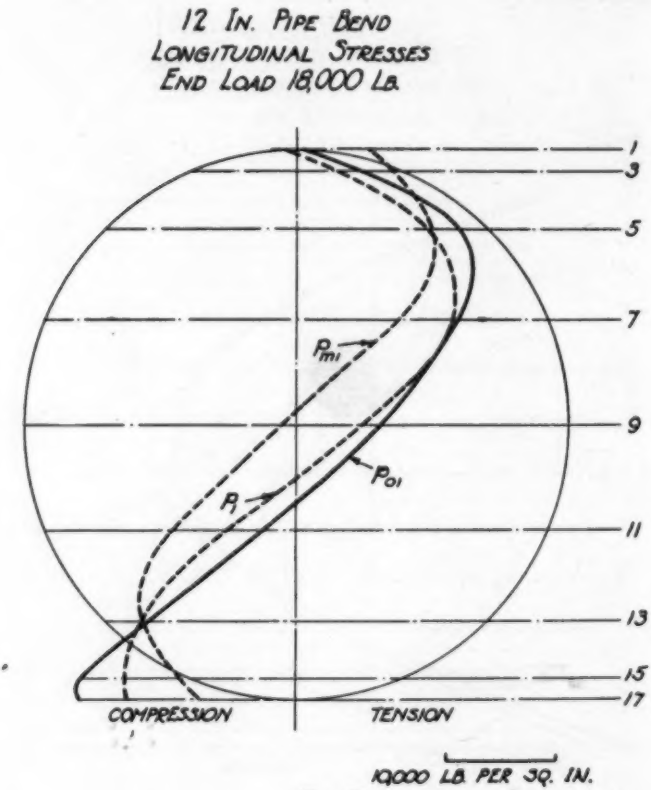


FIG. 11.

or initial load were used. In this way it was expected that the permanent strains, if any, would be eliminated, but, as explained below, this was not quite the case.

Fig. 19 shows curves for the permanent longitudinal strains

in the 6-in. pipe immediately after the application of the various loads. The strains were obtained as the difference between the tensometer readings at 500 lbs. taken immediately after those

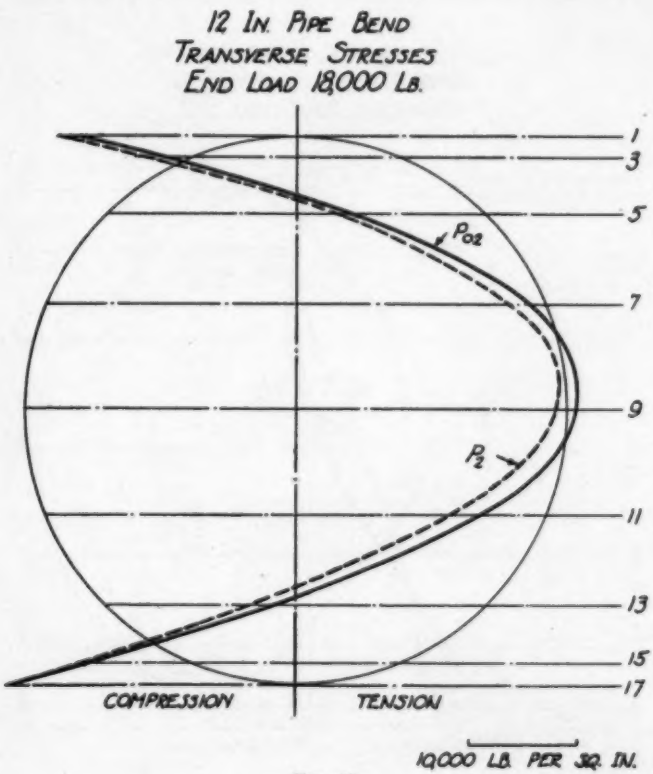


Fig. 12.

loads had been applied and those taken at the original load of 500 lbs.

Fig. 20 gives curves for the total permanent longitudinal strains obtained by subtracting the Berry gage readings at zero load before the tests from those taken at zero load after the tests.

Discussion of the Stress Curves.—On the whole there is a good correspondence in the character of the curves for observed and calculated *longitudinal* stresses, but generally the observed stresses are somewhat greater than the calculated. This peculi-

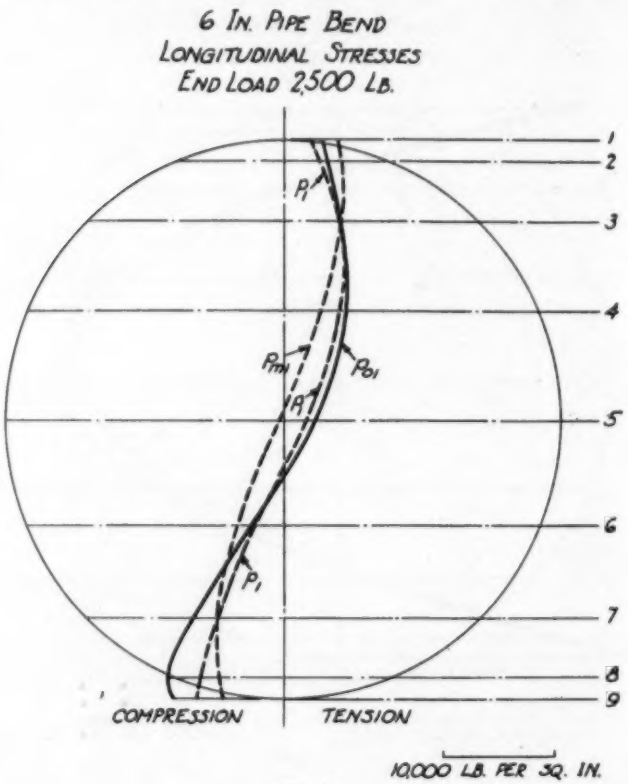


FIG. 13.

arity was observed also in previous tests, in some cases even in a higher degree because the diagrams chosen for illustration referred to loads nearer the limit of elasticity of the pipes.

The curves for the observed and calculated *transverse* stresses show a very close agreement, more so than in any previous tests, perhaps again because they correspond to relatively more moderate loads.

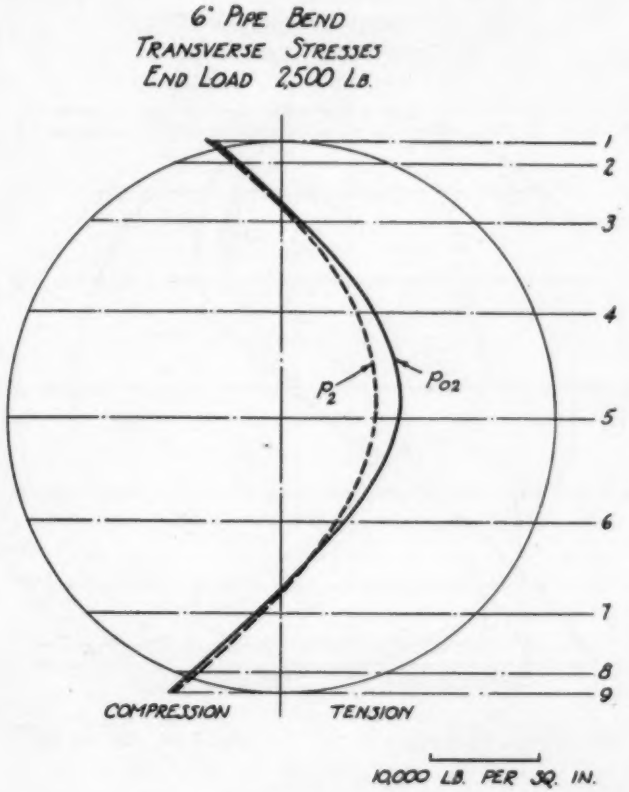


FIG. 14.

In the third paper the discrepancy between the observed and calculated stresses was explained as being due to a flow of the

metal at certain parts of the section, for since the strains were measured relative to the initial load condition, they would comprise both the elastic and plastic deformation.

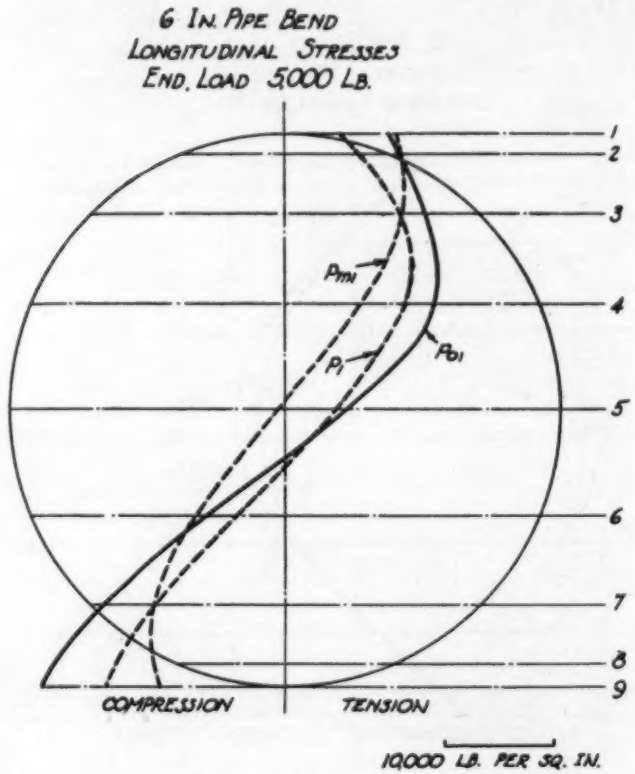


FIG. 15.

In this connection it is of interest to investigate whether the material reached the plastic state in the cases represented on the diagrams where the bends as a whole were within the limit

of proportionality, and for this purpose the magnitude of the "equivalent" stress was determined by the formula:*

$$p_e = \sqrt{p_1^2 - p_1 p_2 + p_2^2} \qquad (12)$$

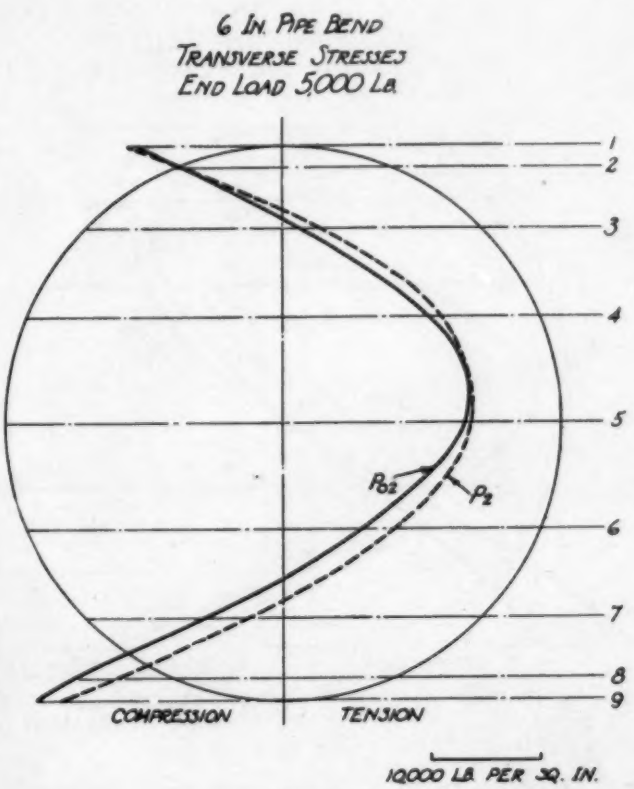


FIG. 16.

Theoretically, the metal should not begin to flow until the value of p_e reaches the yield point, but, applying the formula to

* H. Hencky, ZAMM (Zeitschrift für Angewandte Mathematik und Mechanik), Aug. 1924, p. 331.

the most strained parts in the four pipes here under consideration, it is found that the equivalent stress does not much exceed 20,000 lbs./sq. in. except in the 6-in. pipe under a load of 6500

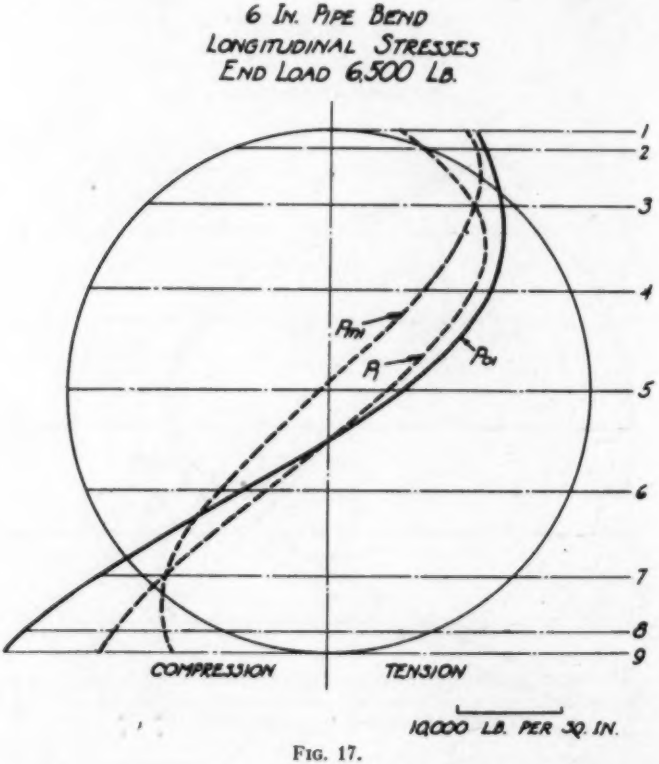


Fig. 17.

lbs., where it is 24,400 lbs. at the lowest point of the section, and even this is probably below the yield point. Hence there should be no flow of the material in any of those cases, but since it did actually occur, the most plausible explanation seems to be that the material was under an initial strain, probably produced in the process of bending.

Accepting as a fact that the plastic state was actually reached locally in certain regions, we can explain the discrepancy between the calculated and observed longitudinal stresses in case of the 8-in., 10-in. and 12-in. pipes, but this explanation fails in

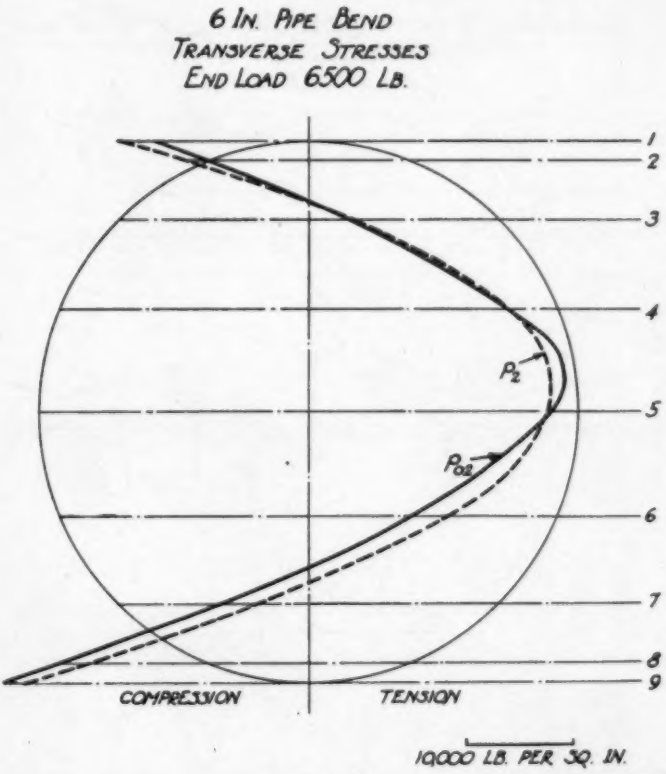


FIG. 18.

case of the 6-in. pipe. Here the permanent strains were supposed to be eliminated by taking the strains relative to the reference load applied immediately after the load in question,

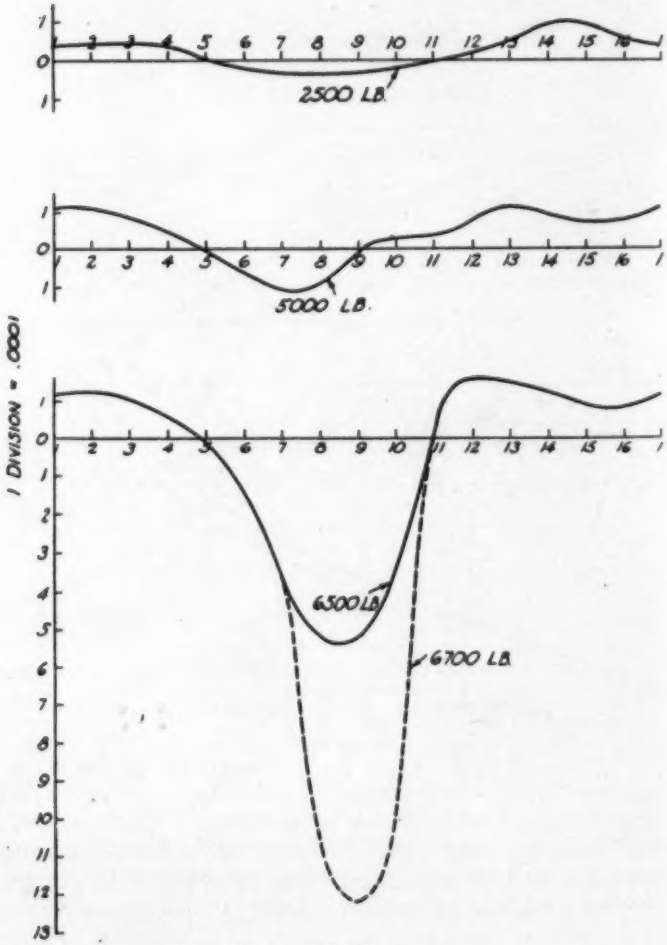

6 IN. PIPE BEND
OBSERVED PERMANENT LONGITUDINAL STRAINS
REMAINING AFTER THE VARIOUS LOADS
2500 LB.
5000 LB.
6500 LB.
6700 LB.
1 DIVISION = .0001

FIG. 19.

and yet the observed longitudinal stress at the top of the section was up to 100 per cent in excess of the calculated stress. A reasonable explanation of this discrepancy seems to be the inter-

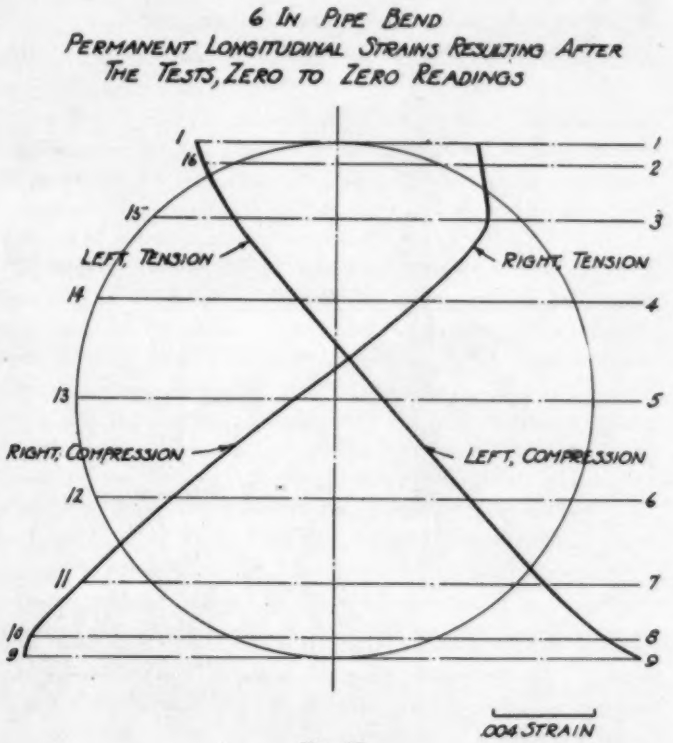


FIG. 20.

action between the plastic parts and the surrounding elastic regions, which latter exert a restraining influence on the former. It must be borne in mind that in the present case incipient plasticity probably does not extend much below the surface of the pipe, because the transverse stresses decrease rapidly towards the middle surface.

Consider now a shallow plastic region, as indicated by the hatched area in Fig. 21, existing, for instance, at the outer surface of the pipe wall at the top of the section, where it

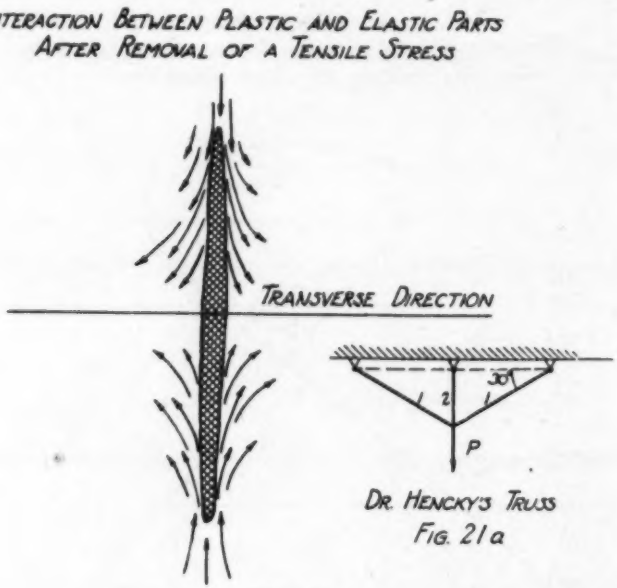


FIG. 21.

would be produced by longitudinal tension and transverse compression. In the plastic part the stress is everywhere at the yield point and in the neighboring elastic material the stress is at the yield point at the boundary of the plastic material and elsewhere more highly stressed than it would be if plasticity did not exist. When the load is removed, the plastic material contracts first until the tension is zero, but due to the pull of the surrounding elastic material it will be somewhat compressed and suffers additional contraction. Hence the measured strain in the plastic region will be greater than the elastic, and in fact

greater than that corresponding to the yield point, although the plastic flow that occurs during the application of the load is eliminated. After this process the material in and surrounding the region of plastic flow is left in a strained condition. If the same load is applied again, the strength of the material will suffer. The state of stress after the load is removed is indicated diagrammatically in Fig. 21.

The action here described is explained in a very elegant and lucid manner by Dr. Hencky in the paper in ZAMM, referred to above, where it is illustrated by considering the loading of a statically indeterminate simple truss consisting of three bars of different sectional area. See Fig. 21a.

The equilibrium which is established in the pipe bend between the bending moment and the stress couple after local plastic flow has occurred under a given load is a result of an adjustment between the elastic and the plastic strains. The plastic parts give way under a constant stress equal to that at the yield point, and the adjacent elastic parts are called upon to take a greater share of the load. So also in Dr. Hencky's truss, bar (2) is first stressed beyond the elastic limit and takes a permanent elongation, while the bars (1) are yet elastic and carry a relatively greater part of the load than when (2) was elastic. If now the load is removed we have a state of stress analogous to that in the pipe illustrated in Fig. 21, where the elastic material adjacent to the plastic region exerts a pull in directions oblique to the plastic elongation.

Table 3 gives a summary of the maximum longitudinal stresses and Table 4 of the corresponding transverse stresses in the four pipes.

An inspection of Table 3 shows that in all cases when the limit of proportionality in the deflection of the bend was reached, the calculated maximum *longitudinal* stress at the middle surface ran from 21,000 to 26,000 lbs./sq. in., the absolute maximum occurring as a compressive stress in the lower part of the section. This confirms the practical rule, given in the third

TABLE 3

Maximum Longitudinal Stresses, Pounds per Square Inch

Description of Bend	8" Bend No. 6		10" Bend No. 7		12" Bend No. 8		6" Bend No. 9		
Load in Pounds	12,000	23,000	18,000	36,000	18,000	30,000	5000	6500	8500
		Limit of Proportionality		Limit of Proportionality		Limit of Proportionality			Limit of Proportionality
Tension: Calc. Stress Middle Surf.	11,500	22,000	10,900	21,800	12,800	21,300	10,800	14,400	18,400
Tension: Calc. Stress Outer Surf.	11,400	21,900	11,400	22,800	14,600	24,300	11,500	15,000	19,600
Tension: Obs. Stress Outer Surf.	15,300				15,900		14,000	16,700	
Compression: Calc. Stress Middle Surf.	13,200	25,200	13,100	26,200	14,500	24,200	12,300	16,000	20,900
Compression: Calc. Stress Outer Surf.	17,800	34,100	17,500	35,000	16,000	26,700	16,500	21,500	28,100
Compression: Obs. Stress Outer Surf.	25,500	Excessive		Excessive	20,000	Excessive	22,100	30,900	Excessive

TABLE 4

Maximum Transverse Stresses, Pounds per Square Inch

Description of Bend	8″ Bend No. 6		10″ Bend No. 7		12″ Bend No. 8		6″ Bend No. 9		
Load in Pounds	12,000	23,000	18,000	36,000	18,000	30,000	5000	6500	8500
		Limit of Proportionality		Limit of Proportionality		Limit of Proportionality			Limit of Proportionality
Tension: Calc. Stress near Neutral Axis	17,500	33,500	18,000	36,000	24,200	40,300	17,100	22,200	29,100
Tension: Obs. Stress near Neutral Axis	15,200				25,500		17,000	23,600	
Compr.: Calc. Stress at Lowest Point	20,800	39,900	21,200	42,400	26,100	43,500	20,300	26,400	34,500
Compr.: Obs. Stress at Lowest Point	19,700				26,500		22,500	28,000	

paper and referred to above, that so long as that stress does not exceed 20,000 lbs./sq. in. in the most strained parts, there will be no appreciable permanent set in the pipe bend. It is assumed that the pipe is made from a good quality of iron or steel, having a yield point of not less than about 25,000 lbs./sq. in.

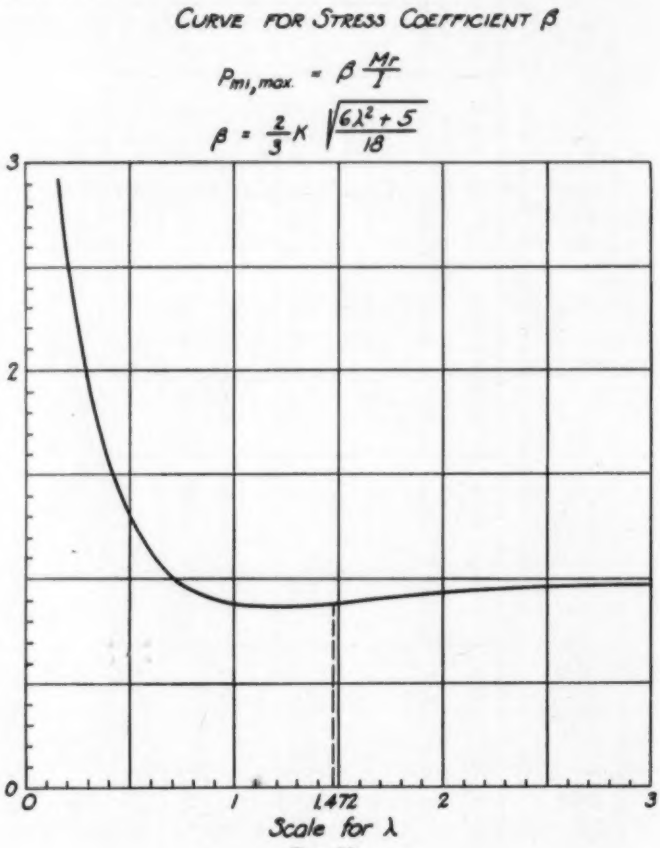


Fig. 22.

The maximum longitudinal stress at the middle surface can be readily calculated from formula (11) or (11′). Apart from p_c, which can be dealt with separately, the expression for $p_{m1,\ max.}$ can be conveniently represented by means of a curve, such as given in Fig. 22. This was suggested by Mr. Arthur Edwards* who proposed to write:

$$p_{m1,\ max.} = \beta \frac{Mr}{I}, \tag{13}$$

where

$$\beta = \frac{2}{3} K \sqrt{\frac{6\lambda^2 + 5}{18}}, \tag{14}$$

and Mr/I is the maximum stress according to the ordinary linear law of bending. When β is plotted on λ as abcissae, it is easy to determine, and by applying it as a factor to the maximum stress calculated by the linear law we obtain $p_{m1,\ max.}$. Usually λ is smaller than or not much greater than unity, but when it is equal to 1.472 the expression under the root sign in (14) becomes equal to one. Since the maximum stress occurs when

$$\sin \theta = \sqrt{\frac{6\lambda^2 + 5}{18}},$$

and since $\sin \theta$ cannot be greater than one, it is seen that for all values of $\lambda \lesseqgtr 1.472$ the maximum value of $p_{m1,\ max.}$ must be found at the top and bottom of the section where $\theta = \pm \frac{\pi}{2}$, and then

$$\beta = \frac{12\lambda^2 - 2}{12\lambda^2 + 1} \tag{14′}$$

It is seen also from Fig. 22, that in such case and, in fact, whenever λ is greater than about .75, β is smaller than unity, so that the maximum stress in the middle surface will fall below what it would be according to the linear law.

Table 4 shows that when the pipe bends are at the limit of

* Of the Stone & Webster Engineering Corporation.

proportionality, the calculated maximum *transverse* stresses are of the order of 35,000 to 45,000 lbs./sq. in. and hence ordinarily beyond the yield point of the material. These stresses have therefore no physical significance and cannot be used as a criterion of strength, but they indicate that a flow of the material has taken place at the outer or inner surface of the pipe wall.

Permanent Strains.—Fig. 19 gives the curves for the permanent strains that remained in the 6-in. pipe bend after the load, stated on the various curves, had been applied and again reduced to 500 lbs. The strains were obtained as the difference between the reading at the latter load of 500 lbs. and that taken at the initial load of 500 lbs.

A permanent strain was found already after application of a load of 2500 lbs., and although of small magnitude, it seemed to follow a well defined sinuous curve with a positive maximum between points 14 and 15, a negative maximum at point 8 and a saddle-like depression at point 1. These features are found more or less consistently in the curves for 5000 lbs. and 6500 lbs. The dotted curve for 6700 lbs. shows that in raising the load from 6500 lbs. to 6700 lbs. a pronounced flow of the metal took place between points 7 and 11. It was at point 7 that the pointer of the tensometer was observed to move after the load of 6500 lbs. had been reached.

Fig. 20 gives the curves for total permanent strains remaining after the tests, obtained as the difference between the Berry gage readings at zero load after the tests and the initial Berry gage readings at zero load before the tests. These strains are plotted as in the stress diagrams, being marked off as abscissae on the distance from the neutral axis 5–13 as ordinates. They show the close approach of the plastic strains to the linear law. The neutral axis, as indicated by zero strains, is seen to lie somewhat above the center of the section.

The Strain Lines.—The strain lines on the 8-in. and 10-in. pipes were very pronounced and essentially of the same pattern as shown in the photograph, Fig. 23, and in the diagram-

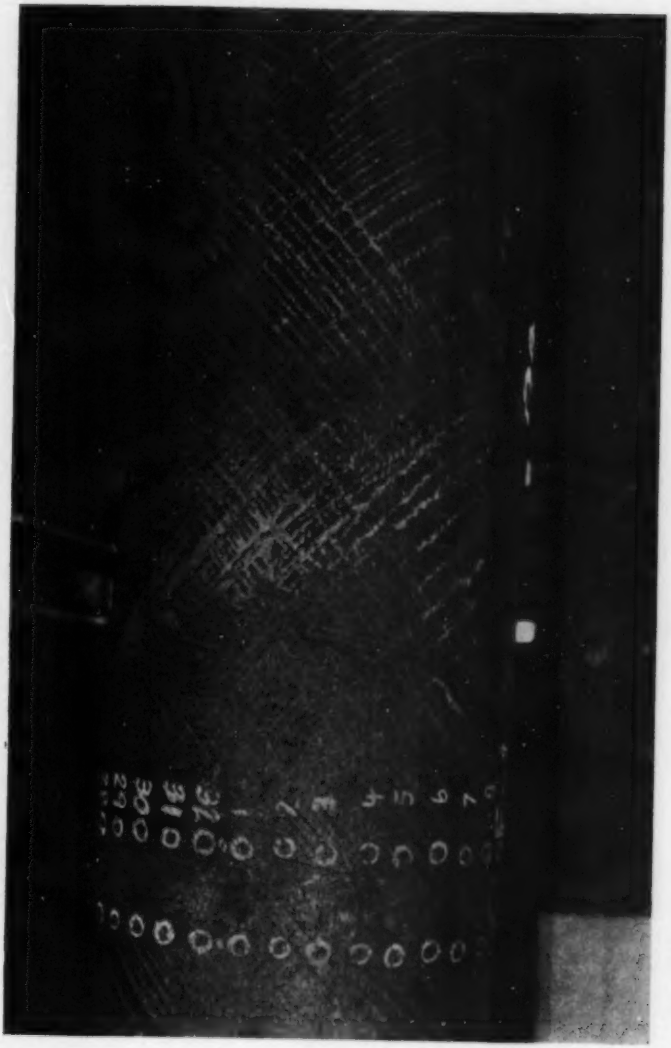


FIG. 23. Eight-Inch Pipe Bend. Strain Lines.

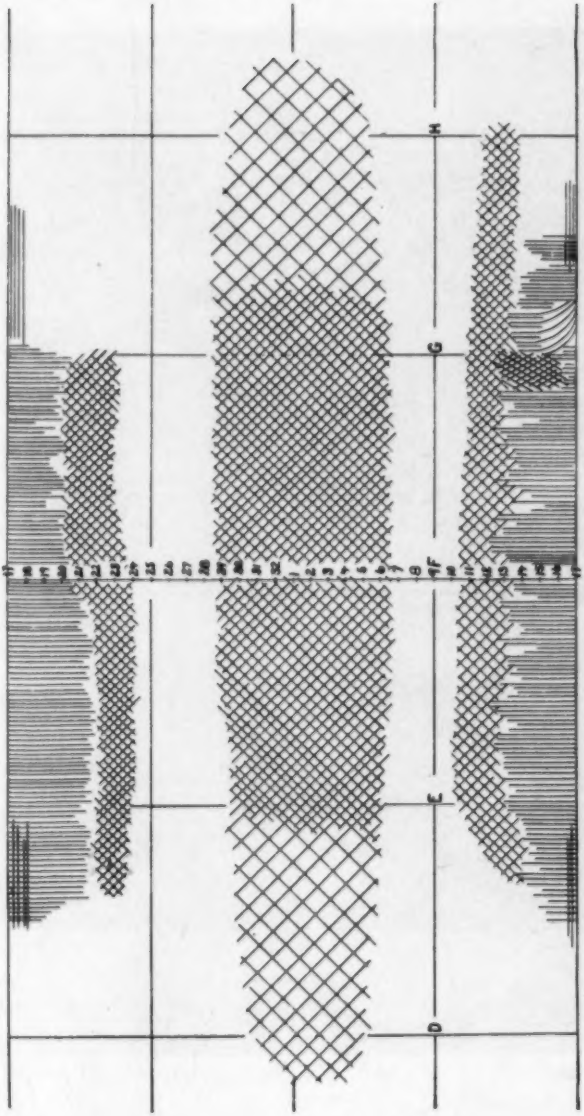


FIG. 24. Eight-Inch Pipe Bend. Development of outer surface showing distribution of strain lines.

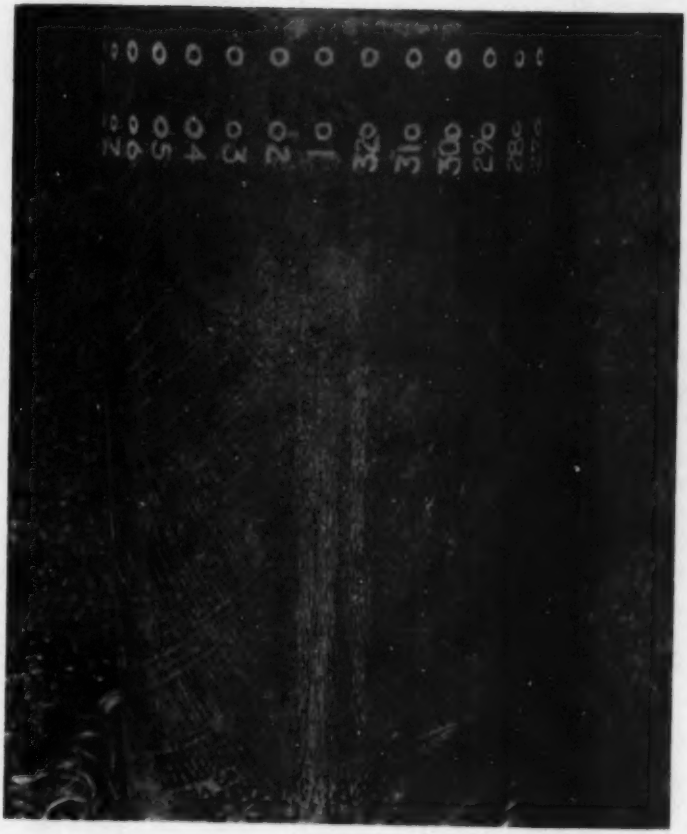


FIG. 25. Twelve-Inch Pipe Bend. Strain Lines.

matic sketch, Fig. 24, both for the 8-in. pipe. The lines run diagonally at about 45 degrees to the axis and are continued in full strength across the top of the pipe. In the 12-in. pipe the lines were likewise very pronounced, but, as seen in Fig. 25, they do not cross the top of the pipe, being here discontinued over an arc of about 20 degrees, and in the intervening belt they are replaced by one or two bunches of wavy longitudinal lines, apparently strain lines.

This difference in the pattern of the lines may perhaps be explained as due to the relatively small wall thickness of the 12-in. pipe, which caused greater flattening of the section and greater transverse stresses at the top than in the heavier 8-in. and 10-in. pipes. Transverse compression at the top, where the longitudinal stresses are small or even compressive—see Figs. 11 and 12—produce shear stresses in planes passing through the longitudinal tangent to the middle surface and inclined at 45 degrees to the tangent plane. These planes intersect the surface of the pipe in longitudinal lines and if the shear stresses are so great as to produce flow of the metal, they may perhaps cause longitudinal strain lines of the type observed on the 12-in. pipe.

Horizontal Deflection of the Ends.—The approach of the two ends of the pipe bend to each other, as determined by direct measurement with steel tape between the end punch marks A and K, was compared with the theoretical value calculated from the formula:

$$\Delta x = \frac{P}{EI}\int Ky^2 ds. \tag{15}$$

As explained in the second paper, y is here the ordinate of any point on the axis of the pipe, taking the line of action of the end forces P as axis of abscissae. The integration extends along the axis from A to K and is performed graphically by plotting the values of Ky^2 as ordinates on the developed girth of the axis of the pipe as abscissae. Table 5 gives the calculated and measured values of Δx for the four pipes corresponding to var-

ious net loads; that is the actual load minus the initial load. The corresponding gross loads are the same as those for which the stress diagrams are worked out, and as given in Tables 3 and 4.

TABLE 5

Horizontal Deflection of the Ends of the Pipe, Δx

	8-in. Pipe No. 6		10-in. Pipe No. 7		12-in. Pipe No. 8		6-in. Pipe No. 9		
Load in Pounds	10500	21500	16000	34000	13000	25000	4500	6000	8000
Δx, Calculated, Inches	.25	.51	.26	.56	.32	.62	.31	.41	.55
Δx, Measured, Inches	.25	.56	.25	.56	.32	.63	.30	.41	.55

There is seen to be a very close correspondence between the calculated and the measured deflections, testifying to the correctness of the formula. This is important, because in practice the deflection is generally what it is desired to know.

Deformation of the Transverse Section.—The flattening of the transverse section was determined by caliper measurement, and one-half the change in vertical diameter gave the measured deflection. This was compared with the calculated deflection at the top point B determined from formula (17′) in the first paper:

$$\Delta y_B = \frac{KMR}{EI} \frac{120r^5}{480h^2R^2 + 101r^4}. \tag{16}$$

The results are given in Table 6.

TABLE 6

Vertical Deflection of the Top Point B of the Section at F, Δy_B

	8-in. Pipe No. 6		10-in. Pipe No. 7		12-in. Pipe No. 8	6-in. Pipe No. 9		
Load in Pounds	10500	21500	16000	34000	13000	4500	6000	8000
Δy_B, Calculated, Inches	.010	.021	.017	.037	.033	.010	.013	.018
Δy_B, Measured, Inches	.010	.028	.016	.039	.034	.009	.012	.018

The correspondence between the calculated and observed values is again very close, and it appears that so long as the horizontal deflection of the pipe as a whole is within the limit of proportionality, formula (16) is very accurate. Although this may not be of direct practical importance, it is of interest because it serves to confirm the theory.

E. The Stresses Due to Steam Pressure

In practice the stresses due to steam pressure are added to those due to bending.

The transverse or hoop stress in a pipe, straight or curved, is given by the formula:

$$p_{2s} = \frac{rp}{t}, \qquad (17)$$

where p is the steam pressure per unit area.

In a pipe bend there will be at the same time a longitudinal or axial stress, which is ordinarily and approximately equal to one-half of the hoop stress. This is not so obvious and is true only under certain conditions and assumptions.

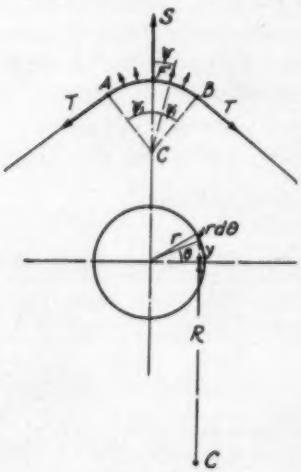


Fig. 26.

Consider a circular pipe bend AB inserted between two straight tangent pipes, forming an angle $2\psi_1$ with each other, equal to the arc covered by the bend. Let S be the resultant of all the steam pressures acting on the bend. See Fig. 26. By symmetry S must pass through the center of the bend and bisect the angle $2\psi_1$; its magnitude is:

$$S = 2\int_0^{2\pi}\int_0^{\psi_1} rd\theta(R + y)d\psi \cdot p \sin\theta \cos\psi = 2\pi r^2 p \sin\psi_1, \quad (18)$$

bearing in mind that:

$$y = r\sin\theta, \int_0^{2\pi} \sin\theta d\theta = 0 \qquad \text{and} \qquad \int_0^{2\pi} \sin^2\theta d\theta = \pi.$$

Now it is assumed that the straight pipes are so long and flexible that they do not offer any appreciable resistance to transverse deflection or to bending. In that case there will be no shear forces and no couples acting at A and B due to the steam pressures, and S is balanced entirely by two forces T acting along the axes of the straight pipes. Thus:

$$S = 2T\sin\psi_1,$$

$$\therefore \quad T = \frac{2\pi r^2 p \sin\psi_1}{2\sin\psi_1} = \pi r^2 p. \quad (19)$$

When the bend forms a semi-circle $\psi_1 = \pi/2$ and $T = S/2$, which is otherwise obvious. For a right-angled bend, $\psi_1 = 45°$ and $T = S/\sqrt{2}$, but (19) gives in all cases the same value for T. In the limit, when ψ_1 is very small, S approaches zero, but theoretically T should retain the same value. Actually in that case S is taken up partly by the small shear forces caused by the resistance to bending of the straight pipes, and the accurate value of T, which is then very small, can not be found except by taking into account the deflection of the pipes at A and B.

Under the assumption that the straight pipes are absolutely flexible the case is similar to that of a flexible string on a smooth curve. The tensile force acting normal to any cross section is

everywhere equal to T and there are no shear forces or bending moments in the bend. Hence the longitudinal stress due to the steam pressure, when the angle of the bend is considerable, is everywhere approximately:

$$p_{1a} = \frac{\pi r^2 p}{2\pi rt} = \frac{rp}{2t}, \tag{20}$$

or one-half of the hoop stress. This longitudinal stress, being tensile, will augment the stress in the upper part of the section at F and reduce it in the lower part. As determined by (20) it must be considered generally an outside limit to the true value, since in practice the straight pipes will never be completely flexible. In fact, if the straight pipes are short, the formula is inapplicable, but in such case the longitudinal stress due to steam pressure will generally be negligible.

F. Alternative Theory

The analysis in the third as well as in the present paper was carried out on the basis of the formulas for the stresses (5), (6), (7) and (8) given above, and the underlying assumption was that the strains denoted by ϵ_{01}, ϵ_{02}, ϵ_1 and ϵ_2 were all "combined" strains, that is, each of them resulted from the combined or superposed effect of longitudinal and transverse stresses. This assumption is certainly correct for ϵ_{01}, the observed longitudinal strain, which was measured directly on the surface of the pipe by strainmeter, but is not as obviously correct for the other strains. Since opinions may differ on this point, we shall examine whether the general conclusions, at which we have so far arrived, will be valid if we make the assumption, which to some may appear more plausible, that ϵ_{02}, ϵ_1 and ϵ_2 are actually "simple" strains, such as would be produced under the action of of single stresses respectively p_{02}, p_1, and p_2, each one acting alone. We denote these simple strains by $\bar{\epsilon}_{02}$, $\bar{\epsilon}_1$ and $\bar{\epsilon}_2$ and have then

$$p_{02} = E\bar{\epsilon}_{02}, \qquad p_1 = E\bar{\epsilon}_1, \qquad p_2 = E\bar{\epsilon}_2. \tag{21}$$

As explained in the first paper, the calculated longitudinal strain was found on the assumption that there were no strains in the neutral plane of the pipe section and that plane transverse sections remained plane in bending even when flattening of the section occurred. At every point the strain was assumed to be the same throughout the wall thickness. We now propose to consider the strain so determined as a simple strain $\bar{\epsilon}_1$ which does not comprise the strain $-\bar{\epsilon}_2/m$ due to transverse stress, and we assume that the latter strain, where it exists, superposes itself on the longitudinal strain. In accordance with formula (8) in the third paper we write:

$$\bar{\epsilon}_1 = \frac{r}{R}\left(\sin\theta - \frac{6r^4\sin^3\theta}{24h^2R^2 + 5r^4}\right)\frac{\Delta d\psi}{d\psi}. \tag{22}$$

The transverse strains ϵ_{02} and ϵ_2 were both calculated under the assumption that the middle surface of the pipe wall is inextensible; ϵ_{02} by determining the curvature on the basis of direct measurement of the various diameters; ϵ_2 by a purely theoretical calculation of the curvature. It is now proposed to regard both of these strains as fundamental or simple, and in accordance with formula (4) we write for the outer surface:

$$\bar{\epsilon}_{02} = h\left(\frac{1}{r'} - \frac{1}{r}\right), \tag{23}$$

and from (16) of the third paper:

$$\bar{\epsilon}_2 = \frac{18hr^3\cos 2\theta}{24h^2R^2 + 5r^4}\frac{\Delta d\psi}{d\psi}. \tag{24}$$

In order to find $\Delta d\psi/d\psi$ we substitute the value of $p_1 = E\bar{\epsilon}_1$ in the expression for the bending moment

$$M = \int_0^{2\pi}\int_{-h}^{+h} r p_1 y\, dz\, d\theta, \tag{25}$$

as explained in the third paper, and get

$$M = \frac{2\pi hr^3E}{R}\,\frac{48h^2R^2 + r^4}{48h^2R^2 + 10r^4}\,\frac{\Delta d\psi}{d\psi}$$

and, with $2\pi hr^3 = I$, we find

$$\frac{\Delta d\psi}{d\psi} = \frac{KMR}{EI}. \tag{26}$$

Hence from (22):

$$p_1 = E\bar{\epsilon}_1 = \frac{KMr}{I}\left(\sin\theta - \frac{6r^4\sin^3\theta}{24h^2R^2 + 5r^4}\right), \tag{27}$$

which is the same as formula (10) given above for the stress at the middle surface, p_{m1}, but which according to this theory exists throughout the wall thickness.

From (24):

$$p_2 = E\bar{\epsilon}_2 = \frac{KMr}{I}\frac{18hr^2R\cos 2\theta}{24h^2R^2 + 5r^4}. \tag{28}$$

This stress is represented by a curve which is symmetrical with respect to the neutral axis, and which intersects that axis at the same point as the curve for p_2 determined by (8). The maximum values of the stresses as given by these curves are very nearly the same.

In order to determine p_{01}, the observed longitudinal strain at the surface, we must go back to the fundamental formulas of elasticity.

The combined strains are expressed in terms of the simple strains by the formulas

$$\epsilon_{01} = \bar{\epsilon}_{01} - \frac{\bar{\epsilon}_{02}}{m}, \tag{29}$$

$$\epsilon_{02} = \bar{\epsilon}_{02} - \frac{\bar{\epsilon}_{01}}{m}, \tag{30}$$

where $\bar{\epsilon}_{03}$ is neglected since p_{03}, the stress normal to the pipe wall, is assumed to be zero. In these equations we know ϵ_{01}, obtained by direct measurement, and $\bar{\epsilon}_{02}$ which is calculated on the basis of the measured diameters. Hence the two unknown strains are,

$$\bar{\epsilon}_{01} = \epsilon_{01} + \frac{\bar{\epsilon}_{02}}{m}, \tag{31}$$

$$\epsilon_{02} = \frac{m^2 - 1}{m^2}\bar{\epsilon}_{02} - \frac{\epsilon_{01}}{m}. \tag{32}$$

Substituting these values in equations (5) and (6):

$$p_{01} = \frac{m^2 E}{m^2 - 1}\left(\epsilon_{01} + \frac{\epsilon_{02}}{m}\right) = \frac{m^2 E}{m^2 - 1}\left(\epsilon_{01} + \frac{m^2 - 1}{m^3}\bar{\epsilon}_{02} - \frac{\epsilon_{01}}{m^2}\right)$$

$$= E\left(\epsilon_{01} + \frac{\bar{\epsilon}_{02}}{m}\right), \tag{33}$$

$$p_{02} = \frac{m^2 E}{m^2 - 1}\left(\epsilon_{02} + \frac{\epsilon_{01}}{m}\right) = \frac{m^2 E}{m^2 - 1}\left(\frac{m^2 - 1}{m^2}\bar{\epsilon}_{02} - \frac{\epsilon_{01}}{m} + \frac{\epsilon_{01}}{m}\right)$$

$$= E\bar{\epsilon}_{02}. \tag{34}$$

The last equation could of course be written down directly from (21).

Comparing now the observed longitudinal stress as found by (33) with that found in the foregoing from (5), which latter in the present symbols becomes:

$$p_{01} = \frac{m^2 E}{m^2 - 1}\left(\epsilon_{01} + \frac{\bar{\epsilon}_{02}}{m}\right),$$

they are seen to differ only by the factor $m^2/(m^2 - 1)$. With $m = 10/3$ the value of this factor is 1.099 so that the new formula gives a stress which at all points is 9 per cent smaller than according to previous analysis.

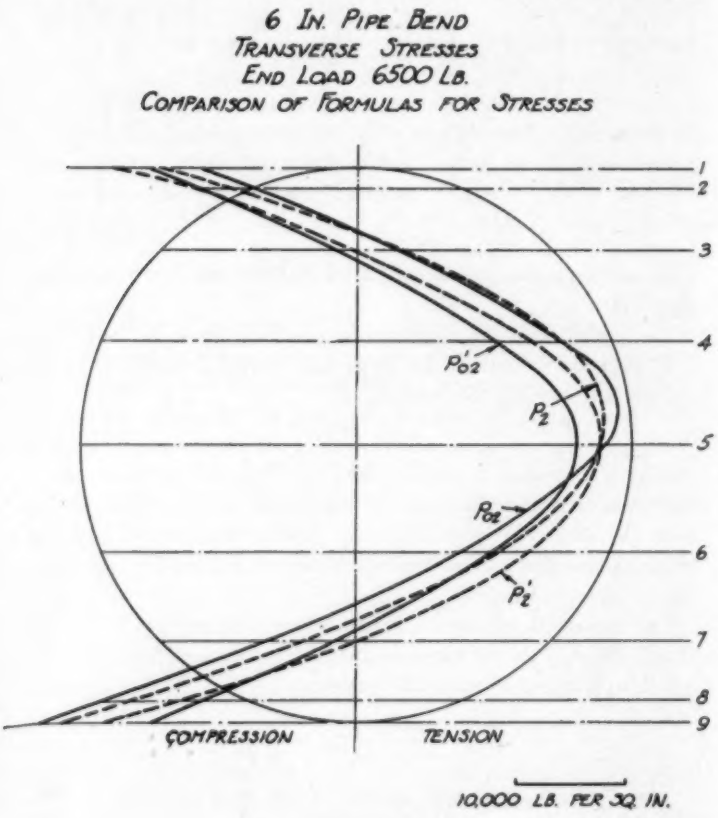
The observed transverse stress determined by (34) is not so simply related to the stress found in previous analysis from (6) which in the present symbols becomes:

$$p_{02} = \frac{m^2 E}{m^2 - 1}\left(\bar{\epsilon}_{02} + \frac{\epsilon_{01}}{m}\right).$$

The effect of using (34) instead of (6), as also of using (28) instead of (8), is a small lowering of the entire stress curve relative to the neutral axis in our previous diagrams but the maximum values do not seem to change much.

In order to illustrate this point, the curves for observed and calculated transverse stresses for the 6-in. pipe bend are given

in Fig. 27 for a load of 6500 lbs. The curves marked p'_2 and p'_{02} are calculated from (28) and (34), and those marked p_2 and p_{02} are calculated by (8) and (6) respectively and are reproduced



p_2 : Formula (8)
p_{02} : Formula (6)
p'_2 Formula (28)
p'_{02} Formula (34)

FIG. 27.

from Fig. 18. This diagram is typical for the differences in the stress curves obtained by the two methods, at least as far as can be judged from the present investigation.

No special diagram is given for the longitudinal stresses. The calculated stress p_1 is already represented by the curves for p_{m1} in the foregoing analysis and the new curves for p_{01} can be obtained from those already given by reducing the stresses by 9 percent.

Summing up, it is seen that the formulas here obtained are much simpler than those used in the third and in the present paper. The resulting stresses do not differ much from those found in previous analysis and the general conclusions at which we have arrived in the foregoing hold good. Notably the practical rule proposed above, that the calculated longitudinal stress at the middle surface, which is the same as p_1 given by (27), can be used as the strength criterion remains valid a fortiori since the longitudinal stress according to the new theory remains unaltered throughout the wall thickness.

The difference between the two theories is somewhat analogous to that which is found in the treatment of plane plates, where one treatment is based on a state of plane stress, another on a state of plane strain. In the previous theory it was assumed that there was no transverse strain but only a transverse stress at the middle surface. According to the present theory this condition is reversed.

G. Summary and Conclusions

The tests here analyzed seem to show that the formulas established in previous papers for pipes of moderate wall thickness hold good also for more thick-walled high-pressure pipes. The theory that the longitudinal stresses in the middle surface can be used as a strength criterion seems to apply to high-pressure pipes as well as to moderate-pressure pipes. Again, it was found that so long as the maximum longitudinal stress, tensile or compressive, at the middle surface does not exceed 20,000 lbs./sq. in.

there will be no appreciable permanent set in a pipe bend. This rule is expected to hold for all pipes designed for moderate or high pressure, whether lap-welded or seamless, of a radius of curvature not less than some four or five diameters, provided also that the material from which the pipes are made is a ductile low-carbon steel or iron of good quality, with a yield point of not less than 25,000 lbs./sq. in.

Although apparently, when the aforesaid stress is reached at the middle surface, the pipe as a whole is within the limit of elasticity, yet it is likely that overstrains will exist locally in certain parts of the pipe and that repeated applications of the same load will in time produce a permanent set. It is recommended, therefore, to design for a substantially lower working stress, say 16,000 lbs./sq. in. This, then, is the stress existing at the middle surface, but allowance should be made for the compressive stress due to end forces, formula (9), and eventually for the tensile stress due to internal steam pressures, as given by formula (20).

The alternative theory for calculating the stresses proposed in Chapter F seems preferable on account of its simplicity, and it will be noticed that the formulas (11) and (13) for the longitudinal stress at the middle surface, to be used as the strength criterion, apply in either case.

The question whether the formulas and the strength criterion here proposed apply to pipe bends of very short radius of curvature is the object of further research now in progress at the Institute.

H. Acknowledgements

The tests on the 8-in. and 10-in. pipe bends were made by Lieutenants A. M. Morgan and R. S. Hatcher, (CC), U. S. N., in preparation for their thesis, presented June, 1929.

The tests on the 12-in. and 6-in. bends were carried out by Professor I. H. Cowdrey and Professor J. H. Zimmerman.

Acknowledgement is due also to Professor E. F. Miller, head of the Department of Mechanical Engineering, and to Professor H. W. Hayward, in charge of the Testing Materials Laboratory, for permission to use the laboratory.